

## THESIS / THÈSE

### MASTER EN SCIENCES MATHÉMATIQUES

#### Une étude des dérivées directionnelles approximées du premier et du second ordre pour des fonctions convexes non différentiables

Stavaux, Monique; Valenduc, Anne

*Award date:*  
1982

*Awarding institution:*  
Université de Namur

[Link to publication](#)

#### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

#### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

FACULTES UNIVERSITAIRES N.D. DE LA PAIX  
NAMUR  
FACULTE DES SCIENCES

DÉPARTEMENT DE MATHÉMATIQUE

ANNÉE ACADÉMIQUE 1981 - 1982

UNE ÉTUDE DES DÉRIVÉES DIRECTIONNELLES APPROXIMÉES  
DU PREMIER ET DU SECOND ORDRE POUR DES FONCTIONS  
CONVEXES NON DIFFÉRENTIABLES.

MÉMOIRE PRÉSENTÉ POUR L'OBTENTION DU  
GRADE DE LICENCIÉES EN SCIENCES  
MATHÉMATIQUES PAR

PROMOTEUR DU MÉMOIRE :

MR. NGUYEN VAN HIEN

MONIQUE STAVAUX

ANNE VALENDUC



Nous tenons à remercier Monsieur  
NGUYEN VAN HIEN d'avoir accepté la direction  
de ce mémoire et Monsieur Jean-Jacques  
STRODIOT pour l'aide et les conseils qu'il  
nous a prodigués tout au long de l'élabora-  
tion de ce travail.

Nous remercions aussi nos parents  
et amis qui nous ont encouragées durant nos  
études.

## TABLE DES MATIERES.

Introduction .....	4
<u>Chapitre I</u> : Notions d'analyse convexe.....	9
<u>Chapitre II</u> : Etude du $\varepsilon$ - sous-différentiel et de sa fonction d'appui.....	25
II.1. : Définition et exemples de $\varepsilon$ -sous-différentiel....	27
a) Définition .....	28
b) Illustration graphique .....	28
II.2. : Caractérisation du $\varepsilon$ -sous-différentiel .....	29
a) Caractérisation par la conjuguée .....	30
b) Caractérisation par la fonction d'appui .....	31
II.3. : Définition et propriétés de $q_{d,\varepsilon}(\lambda)$ .....	36
a) Définition .....	36
b) Comportement aux limites de $q_{d,\varepsilon}(\cdot)$ pour tout $\varepsilon > 0$ .	36
c) Propriétés de $q_{d,\varepsilon}(\cdot)$ et définition de $r_{d,\varepsilon}(\cdot)$ .....	37
d) Comportement à la limite de $r_{d,\varepsilon}(\cdot)$ .....	39
e) Exemples de calculs des fonctions $q_{d,\varepsilon}(\cdot)$ et $r_{d,\varepsilon}(\cdot)$ .	39
II.4. : Comportement local de $f'_\varepsilon(x_0; d)$ .....	48
II.5. : Exemple d'algorithme du premier ordre .....	52
<u>Chapitre III</u> : Etude de la différentiabilité de $V_{d,\varepsilon}(\cdot)$ .....	56
III.1. : Expression de la dérivée directionnelle de $V_{d,\varepsilon}(\cdot)$ dans une direction $\delta$ .....	58
III.2. : Sous-différentiabilité de $V_{d,\varepsilon}(\cdot)$ .....	67
III.3. : Caractérisations de l'ensemble $\Lambda_{d,\varepsilon}(x_0)$ et conséquences sur l'expression du gradient .....	81
1er cas : $\Lambda_{d,\varepsilon}(x_0)$ est non vide et borné .....	81
2me cas : $\Lambda_{d,\varepsilon}(x_0)$ est non vide et non borné .....	87
3me cas : $\Lambda_{d,\varepsilon}(x_0)$ est vide .....	91
<u>Chapitre IV</u> : Comportement à la limite de $f'_\varepsilon(x_0; d, \delta)$ quand $\varepsilon \rightarrow 0$ .....	96

IV.1. : Etude de $\Lambda_d(\epsilon)$ en fonction de $\epsilon$ .....	98
A) Nouvelle description de $\Lambda_d(\epsilon)$ .....	98
A.1) Interpretation de $a_d^*$ .....	98
A.2) Conséquences sur la classification de $\Lambda_d(\epsilon)$ .	101
B) Comportement de $\Lambda_d(\epsilon)$ quand $\epsilon \rightarrow 0$ .....	106
IV.2. : Etude de $f'_\epsilon(x_0; d)$ en fonction de $\epsilon$ .....	113
A) Définition de $\Theta_d(\cdot)$ et $r_{d, \epsilon}(\cdot)$ sur $R$ .....	113
B) Propriétés de $\Theta_d(\cdot)$ .....	114
C) Comportement de $f'_\epsilon(x_0; d)$ quand $\epsilon \rightarrow +\infty$ .....	121
D) Comportement de $f'_\epsilon(x_0; d)$ quand $\epsilon \rightarrow 0^+$ .....	122
IV.3. : Comportement de $f''_\epsilon(x_0; d, \delta)$ comme fonction de $\epsilon$ .	129
A) Etude de $f''_\epsilon(x_0; d, \delta)$ axée sur la valeur de $a_d^*$ ...	129
B) $f''_\epsilon(x_0; d, \delta)$ comme approximation de la dérivée de Dini .....	137
1) Définition .....	137
2) Lien avec la différentiabilité au sens de Mignot .....	138
3) Résultats fondamentaux sur la convergence de $f''_\epsilon(x_0; d, \delta)$ .....	140
<u>Chapitre V</u> : Etude de $f''_\epsilon(x; d, \delta)$ comme fonction de $x$ et de $(d, \delta)$ .....	151
V.1. : Semi-continuité de $f''_\epsilon(x; d, \delta)$ comme fonction de $x$ .....	152
V.2. : Comportement de $f''_\epsilon(x; d, \delta)$ comme fonction de $(d, \delta)$ .....	154
A) $f''_\epsilon(x; d, \delta)$ est positivement homogène en $(d, \delta)$ ..	154
B) Comportement de $f''_\epsilon(x; d, \delta)$ quand $\ d\  \rightarrow \infty$ .....	155
C) Convexité de $f''_\epsilon(x; d, \delta)$ en fonction de $d$ .....	156
D) Continuité de $f''_\epsilon(x; d, \delta)$ en fonction de $d$ .....	163
V.3. : Règles de calculs sur $f''_\epsilon(x; d, \delta)$ .....	165
<u>Annexes</u> .....	168
<u>Bibliographie</u> .....	177

# INTRODUCTION



En optimisation, nombreux sont les problèmes où la fonction à minimiser est convexe mais non nécessairement différentiable. Ainsi, on rencontre fréquemment des fonctions de coût du type  $\|g(x)\|_1$ ,  $\|g(x)\|_\infty$ ,  $\max_i |f_i(x)|$ ... Pour de telles fonctions, en les points où il n'existe pas, le gradient est généralisé par un ensemble de sous-gradients, appelé sous-différentiel.

La condition d'optimalité assurant que  $x_0$  minimise la fonction  $f$  se transpose en l'appartenance de 0 au sous-différentiel de  $f$  en  $x_0$ .

Cependant, la solution du problème n'est accessible numériquement qu'à une précision d'ordre  $\epsilon$ . Le problème revient donc à caractériser un  $\epsilon$ -minimum de  $f$ . Dans cette optique, nous considérons un "élargissement" du concept de sous-différentiel, le  $\epsilon$ -sous-différentiel. Dès lors, la condition nécessaire et suffisante pour que  $x_0$  minimise  $f$  à  $\epsilon$  près sera l'appartenance de 0 au  $\epsilon$ -sous-différentiel de  $f$  en  $x_0$ . Notons que beaucoup d'algorithmes se basent sur cet outil du premier ordre.

De plus, en vue de généraliser des méthodes utilisant la dérivée seconde telles Newton, Quasi-Newton,... au cas non différentiable, de nombreuses recherches tentent de définir un analogue du deuxième ordre. L'exploration théorique d'une de ces voies sera notre centre d'intérêt au cours de ce mémoire.

Celle-ci est légitimée par les propriétés de la fonction d'appui du  $\epsilon$ -sous-différentiel,  $f'_\epsilon(x; d)$ . En effet, la dérivée directionnelle exacte ne possédant pas ces propriétés intéressantes rendait une telle approche impossible.

L'étude du  $\epsilon$ -sous-différentiel a débuté en 1965, elle n'a cessé d'évoluer et est encore un sujet de recherche aujourd'hui. Donnons-en un bref historique.

1965-1970 : définition et premières propriétés du  $\varepsilon$ -sous-différentiel, parallèlement au développement de l'analyse convexe.

1970-1979 : développement des méthodes du premier ordre. Elles définissent  $x_{m+1}$  à partir de  $x_m$  grâce à la connaissance du  $\varepsilon_m$ -sous-différentiel de  $f$  en  $x_m$ .

1979-1981 : durant ces trois années, l'étude de la dérivée directionnelle approximée du premier ordre fut d'un grand intérêt.

Dans ce mémoire, nous nous basons sur les travaux de cette dernière période dus en majeure partie à J-B Hiriart-Urruty. Nous les développerons et illustrerons par de nombreux exemples afin d'en mieux percevoir la portée.

Nous rappellerons, dans un premier chapitre, des notions de base d'analyse convexe telles : sous-gradient, dérivée directionnelle...

Celles-ci nous permettront d'aborder le concept de  $\varepsilon$ -sous-différentiel dans le second chapitre. Nous étudierons également sa fonction d'appui  $f'_\varepsilon(x; d)$  et développerons un algorithme du premier ordre pour minimiser une fonction convexe non différentiable à  $\varepsilon$  près. Contrairement à la dérivée directionnelle exacte, la fonction  $x \rightarrow f'_\varepsilon(x; d)$ , encore notée  $v_{d, \varepsilon}$ , est localement lipschitz. Munie de cette propriété, elle admet un gradient généralisé au sens de Clarke dont l'expression sera développée dans le troisième chapitre. Nous prouverons aussi que, en tout point  $x$  où elle n'est pas différentiable, la fonction  $v_{d, \varepsilon}(\cdot)$  admet, pour chaque direction, une dérivée directionnelle encore notée  $f''_\varepsilon(x; d, \delta)$ . Dans le cas où  $v_{d, \varepsilon}$  est différentiable, nous donnerons une expression précise de son gradient.

Dans le chapitre suivant, le nouveau concept  $f''_\varepsilon(x; d, \delta)$  sera envisagé comme approximation de la dérivée directionnelle du second ordre de  $f$  en  $x$ , même si cette dernière n'existe pas. Dans ce sens, un des résultats importants est la convergence de  $f''_\varepsilon(x; d, \delta)$  vers la dérivée de Dini quand  $\varepsilon$  tend vers 0.



Afin de compléter l'étude de  $f'_\varepsilon(x; d, \delta)$ , nous l'examinerons au cinquième chapitre comme fonction de  $x$  et de  $(d, \delta)$ . Nous montrerons qu'elle est semi-continue inférieurement en la variable  $x$  et que  $f'_\varepsilon(x; d, d)$  n'est ni convexe, ni continue en la variable  $d$ .

Ce travail ne clôture certainement pas les connaissances sur la sous-différentiabilité à  $\varepsilon$  près, mais suggère des pistes de recherches dotées d'un avenir prometteur.

# notations

$$\cdot \partial f(x_0) = \{x^* \in \mathbb{R}^n : f(x) \geq f(x_0) + \langle x - x_0, x^* \rangle, \forall x \in \mathbb{R}^n\}$$

$$\cdot q_d(\lambda) = \frac{f(x_0 + \lambda d) - f(x_0)}{\lambda}$$

$$\cdot f'(x_0; d) = \inf_{\lambda > 0} q_d(\lambda)$$

$$\cdot \partial_\varepsilon f(x_0) = \{x^* \in \mathbb{R}^n : f(x) \geq f(x_0) + \langle x - x_0, x^* \rangle - \varepsilon, \forall x \in \mathbb{R}^n\}$$

$$\cdot q_{d,\varepsilon}(\lambda) = \frac{f(x_0 + \lambda d) - f(x_0) + \varepsilon}{\lambda}$$

$$\cdot f'_\varepsilon(x_0; d) = \inf_{\lambda > 0} q_{d,\varepsilon}(\lambda)$$

$$\cdot f_d(\lambda) = \sup_{\lambda > 0} q_d(\lambda) = \lim_{\lambda \rightarrow \infty} q_d(\lambda)$$

$$\cdot \pi_{d,\varepsilon}(\mu) = \mu \left[ f(x_0 + \frac{d}{\mu}) - f(x_0) + \varepsilon \right]$$

$$\cdot \Lambda_{d,\varepsilon}(x_0) = \{ \lambda \in \mathbb{R}_0^+ \mid q_{d,\varepsilon}(\lambda) = f'_\varepsilon(x_0; d) \}$$

$$\begin{aligned} \cdot \Pi_{d,\varepsilon}(x_0) &= \{ \mu \in \mathbb{R}^+ \mid \mu = 1/\lambda \text{ for } \lambda \in \Lambda_{d,\varepsilon}(x_0) \} \\ &= \{ \mu \in \mathbb{R}^+ \mid \pi_{d,\varepsilon}(\mu) = f'_\varepsilon(x_0; d) \} \end{aligned}$$

$$\cdot V_{d,\varepsilon}(x) = f'_\varepsilon(x; d)$$

$$\cdot \Theta_d(\varepsilon) = f'_\varepsilon(x_0; d)$$

$$\cdot \partial_{\varepsilon} f(x_0)_d = \{x^* \in \partial_{\varepsilon} f(x_0) : \langle x^*, d \rangle = f'_{\varepsilon}(x_0; d)\}$$

$$\begin{aligned} \cdot f''_{\varepsilon}(x_0; d, \delta) &= \lim_{\delta \rightarrow 0} \frac{f'_{\varepsilon}(x_0 + \delta d; d) - f'_{\varepsilon}(x_0; d)}{\delta} \\ &= \min_{\mu \in \Pi_{d, \varepsilon}(x_0)} \max_{x^* \in \partial_{\varepsilon} f(x_0)_d} \left\{ \mu[\langle x^*, d \rangle - f'_{\varepsilon}(x_0; d)] \right\} . \end{aligned}$$

$$\cdot f''_{\varepsilon}(x_0; d, d) = \min_{\mu \in \Pi_{d, \varepsilon}(x_0)} \left\{ \mu[f'_{\varepsilon}(x_0; d) - f'_{\varepsilon}(x_0; d)] \right\}$$

$$\cdot \alpha_d^* = \text{le sup\'erieur des } \alpha \in \mathbb{R}_0^+ \text{ tq}$$

$$\frac{f(x_0 + \lambda d) - f(x_0)}{\lambda} = f'_{\varepsilon}(x_0; d) \quad \forall \lambda \in ]0, \alpha].$$

$$\cdot \varepsilon_d^* = \text{le sup\'erieur des } \varepsilon \text{ tq}$$

$$\mathcal{Q}_{d, \varepsilon}(\mathbb{R}_0^+) \cap ]-\infty, f_{\infty}(d)] \neq \emptyset.$$


---

# CHAPITRE I

*Notions d'analyse convexe.*

Tout au long de ce mémoire, nous considérons des fonctions convexes, propres, définies sur  $\mathbb{R}^n$  et à valeurs réelles.

Les concepts clés utiles à la résolution du problème d'optimisation : "minimiser une telle fonction non nécessairement différentiable", sont par exemple sous-différentiel, dérivée directionnelle...

C'est pourquoi nous présentons dans ce chapitre différentes notions d'analyse convexe, sans prétention toutefois d'en faire une étude complète.

Nous rappelons d'abord quelques notions élémentaires sur les ensembles et fonctions convexes, sur les fonctions conjuguées, ensuite sur le sous-différentiel et la dérivée directionnelle d'une fonction convexe. Nous généraliserons ces résultats aux fonctions localement lipschitziens pour définir le gradient généralisé.

# 1. Ensembles et fonctions convexes.

## Définitions.

### 1.1. Ensemble convexe.

Soit  $C \subset \mathbb{R}^n$ .

$C$  est un ensemble convexe si  $\sum_{i=1}^m \lambda_i x_i \in C$  où  $\sum_{i=1}^m \lambda_i = 1$  et  $x_i \in C, \forall i, \lambda_i \geq 0$ .

### 1.2. Enveloppe convexe.

L'enveloppe convexe de  $C$ , notée  $co C$  est l'ensemble  $\left\{ \sum_{i=1}^m \lambda_i x_i \mid m \in \mathbb{N}, \sum_{i=1}^m \lambda_i = 1, \lambda_i \geq 0 \text{ et } x_i \in C \text{ pour } 1 \leq i \leq m \right\}$

### 1.3. Fonction convexe.

La fonction  $F$  est convexe sur  $C$  si  $\forall x_1, x_2 \in C$ ,  
 $F(\lambda x_1 + (1-\lambda)x_2) \leq \lambda F(x_1) + (1-\lambda)F(x_2) \quad \forall \lambda \in [0, 1]$ .

### 1.4. Fonction indicatrice.

La fonction indicatrice de  $C$ , notée  $\chi_C$  est telle que :

$$\chi_C(x) = \begin{cases} 0 & \text{si } x \in C. \\ +\infty & \text{sinon.} \end{cases}$$



Remarquons que

L'ensemble  $C$  est convexe si  $\pi_C$  est convexe.

### 1.5 Fonction propre

Une fonction convexe  $F : C \rightarrow \mathbb{R}$  est propre si  $F(x) \neq -\infty \quad \forall x \in C$  et  $F \neq +\infty$ .

### 1.6 Épigraphe de $F$

$$\text{Épi } F = \{ (x, a) \in C \times \mathbb{R} \mid F(x) \leq a \}$$

Remarquons que

Épi  $F$  est convexe si  $F$  est convexe.

### 1.7. Fonction de récession. ([7], p.66)

Si  $F$  est une fonction convexe propre, sa fonction de récession, notée  $F_0^+$  est définie pour chaque vecteur  $y$  comme

$$F_0^+(y) = \sup \{ F(x+y) - F(x) \mid x \in \text{dom } F \}.$$

Cette fonction est convexe, propre et positivement homogène.

De plus, si  $F$  est semi continue inférieurement (voir déf. ci-après).

le théorème 8.5 ([7], p.66) nous enseigne que  $F_0^+$  est aussi sci et pour chaque  $x \in \text{dom } F$ , on a :

$$F_0^+(y) = \sup_{\lambda > 0} \frac{F(x + \lambda y) - F(x)}{\lambda} = \lim_{\lambda \rightarrow \infty} \frac{F(x + \lambda y) - F(x)}{\lambda}.$$



## 2. Continuité des fonctions convexes.

### Définitions.

#### 2.1. Semi-continuité inférieure (s.c.i.)

La fonction  $F: \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$  est s.c.i. sur  $\mathbb{R}^n$  si l'une des deux équivalences est satisfaite:

•  $\forall a \in \mathbb{R}, \{x \in \mathbb{R}^n \mid F(x) \leq a\}$  est fermé

•  $\forall \bar{x} \in \mathbb{R}^n, \liminf_{x \rightarrow \bar{x}} F(x) \geq F(\bar{x})$

où  $\liminf_{x \rightarrow \bar{x}} F(x)$  est la plus petite valeur d'adhérence de l'ensemble  $\{F(x) \mid x \rightarrow \bar{x}\}$ .

#### 2.2. Fonction localement lipschitz (L.L.)

La fonction  $F: C \rightarrow \mathbb{R}$  est L.L. en  $x \in C$  si il existe un voisinage  $N$  de  $x$  et une constante  $K$  tels que  
 $\forall y, z \in N$ , on a :

$$|F(y) - F(z)| \leq K \|y - z\|.$$

### Propositions.

2.3.  $F: \mathbb{R}^n \rightarrow \mathbb{R}$  est s.c.i. si épi  $F$  est fermé.

2.4. Toute fonction convexe propre sur un espace de dimension finie est continue sur son domaine de définition.

2.5. Toute fonction convexe sur un espace de dimension finie est localement lipschitz.

### Définitions.

2.6. Fonction positivement homogène.

La fonction  $F : \mathbb{R}^n \rightarrow \mathbb{R}$  est positivement homogène si  
 $\forall \lambda > 0 \quad F(\lambda x) = \lambda F(x), \quad \forall x \in \mathbb{R}^n.$

2.7. Fermeture de  $F$  ( $Cl F$ ).

$Cl F$  est la plus grande fonction s.e.i. majorée par  $F$ :  
 $\text{épi}(Cl F) = \overline{\text{épi}(F)}$

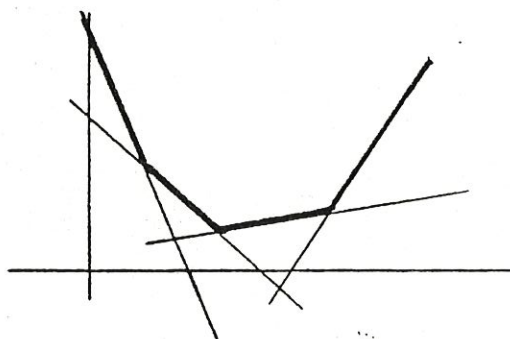
Remarquons que

Si  $F$  est s.e.i., alors  $Cl F = F$

### Proposition

2.8. Toute fonction convexe propre de  $\mathbb{R}^n$  dans  $\mathbb{R}$  est enveloppe supérieure d'une famille de fonctions affines continues.

$$F(x) = \max \{ a_i x + b_i, \quad a_i, b_i \in \mathbb{R}, i \in I \}$$



Dans le cas où  $I$  est fini,  $F$  est polyédrale.

### 3. Fonctions conjuguées.

#### 3.1. Définition du produit scalaire.

Dans le cas d'un espace vectoriel quelconque  $V$ , on note  $V^*$  l'espace vectoriel des formes linéaires continues sur  $V$ , (espace dual) et on définit le produit scalaire externe comme la forme bilinéaire :

$$\langle \cdot, \cdot \rangle : V \times V^* \longrightarrow \mathbb{R}$$

$$(u, u^*) \longmapsto \langle u, u^* \rangle.$$

Comme les fonctions qui nous intéressent sont définies sur  $\mathbb{R}^n$  dont le dual est encore  $\mathbb{R}^n$ , le produit scalaire externe est identique au produit scalaire interne :

$$(\cdot, \cdot) : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$$

Cependant nous garderons la notation  $\langle \cdot, \cdot \rangle$  pour le produit scalaire et des  $*$  pour désigner les éléments du dual.

### 3.2. Définition de la fonction conjuguée

Soit  $F : \mathbb{R}^n \rightarrow \mathbb{R}$ .

La fonction conjuguée (ou polaire) de  $F$ ,  $F^* : \mathbb{R}^n \rightarrow \mathbb{R}$  est définie pour tout  $x^* \in \mathbb{R}^n$  par

$$F^*(x^*) = \sup_{x \in \mathbb{R}^n} \{ \langle x, x^* \rangle - F(x) \}$$

Notons que  $F^*$  est l'enveloppe supérieure de la famille des fonctions affines  $(\langle x, \cdot \rangle - F(x))_{x \in \mathbb{R}^n}$ .

### Propriétés de la fonction conjuguée.

3.3. La conjuguée  $F^*$  est convexe et s.e.i, et ce quelque soit  $F$ .

3.4.  $\forall \alpha \in \mathbb{R}, (F + \alpha)^* = F^* - \alpha$ .

### 3.5. Définition de la bipolaire.

La bipolaire de  $F$ , notée  $F^{**} : \mathbb{R}^n \rightarrow \mathbb{R}$  est définie pour tout  $x \in \mathbb{R}^n$  par

$$F^{**}(x) = \sup_{x^* \in \mathbb{R}^n} \{ \langle x, x^* \rangle - F^*(x^*) \}$$

### Proposition.

3.6. Si  $F$  est convexe, propre, de  $\mathbb{R}^n$  dans  $\mathbb{R}$   
Alors  $F^{**} = F$ .



### 3.7. Définition de la fonction d'appui de $C$ .

Soit  $\chi_C$ , la fonction indicatrice de  $C$  de  $\mathbb{R}^n$ .

La conjuguée de  $\chi_C$ ,  $\chi_C^*$  telle que  $\chi_C^*(x^*) = \sup_{x \in C} \langle x, x^* \rangle$  est appelée fonction d'appui (ou fonction support) de  $C$ .

### Proposition.

3.8. La fonction d'appui d'un ensemble est convexe, s.e.i. et positivement homogène sur  $\mathbb{R}^n$ .

## 4. Sous-différentiabilité des fonctions convexes.

### 4.1 Définition de la sous-différentiabilité d'une fonction

La fonction  $F: \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$  est sous-différentiable en  $x \in \mathbb{R}^n$  si il existe  $x^* \in \mathbb{R}^n$  tel que

$$F(y) \geq F(x) + \langle y - x, x^* \rangle, \text{ pour tout } y \in \mathbb{R}^n.$$

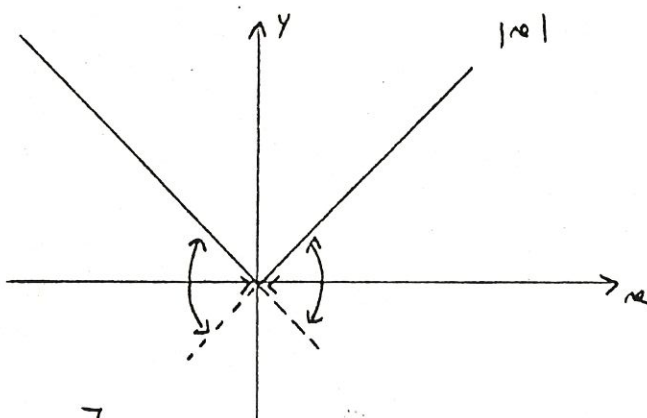
Le point  $x^*$  est un sous-gradient de  $F$  en  $x$ .

L'ensemble de ces  $x^*$  est appelé sous-différentiel et noté  $\partial F(x)$ .

### 4.2. Exemple.

Soit  $f: \mathbb{R} \rightarrow \mathbb{R}$

$$x \mapsto |x|.$$



$$\partial f(0) = [-1, 1]$$

$\mathcal{C}$  est l'ensemble des pentes des fonctions affines minorant  $f$  en 0

### Propositions.

4.3.  $F(x) = \min_{y \in \mathbb{R}^n} F(y) \quad \text{si} \quad 0 \in \partial F(x)$

4.4.  $\partial F(x)$  est un ensemble convexe et compact, pour tout  $x \in \mathbb{R}^n$

4.5.  $\partial F(\cdot)$  est une multifonction s.e.i.

4.6. Soit  $F^*$  la conjuguée de  $F$ . Alors,

.  $x^* \in \partial F(x) \Leftrightarrow F(x) + F^*(x^*) = \langle x, x^* \rangle$

.  $x^* \in \partial F(x) \Leftrightarrow x \in \partial F^*(x^*)$

4.7. Une fonction  $F$  convexe, s.e.i., propre définie sur un espace normé complet est différentiable "presque partout" dans  $\text{dom } F$ , i.e.  $\partial F(x)$  est réduit à un singleton partout sauf sur un ensemble de mesure nulle.

### 4.8. Définition de la dérivée directionnelle.

La dérivée directionnelle de  $F$  au point  $x$  et dans la direction  $d$  est notée  $F'(x; d)$  et

$$F'(x; d) = \inf_{\lambda > 0} \frac{F(x + \lambda d) - F(x)}{\lambda}$$

quand cet infimum existe.

### Propositions.

Considérons  $F$  convexe, une direction  $d$  et un point  $x$  fixés.

4.9. Le quotient  $\frac{F(x + \lambda d) - F(x)}{\lambda}$  est croissant en  $\lambda$  et dès lors

$$F'(x; d) = \inf_{\lambda > 0} \frac{F(x + \lambda d) - F(x)}{\lambda} = \lim_{\lambda \rightarrow 0} \frac{F(x + \lambda d) - F(x)}{\lambda}$$

existe toujours.

### Remarque

Nous notons le quotient  $\frac{F(x + \lambda d) - F(x)}{\lambda}$  par  $q_d(\lambda)$ . Nous avons donc

$$F'(x; d) = \lim_{\lambda \rightarrow 0} q_d(\lambda).$$

4.10.  $\partial F(x) \neq \emptyset$  et pour chaque  $d \in \mathbb{R}^n$ ,

$$F'(x; d) = \max \{ \langle d, x^* \rangle \mid x^* \in \partial F(x) \}$$

On peut dès lors exprimer  $F'(x; \cdot)$  comme la fonction d'appui du sous-différentiel de  $F$  en  $x$ . ( $\delta^*(\cdot \mid \partial F(x))$ )

4.11. Condition d'optimalité.

si  $F(x) = \min_{y \in \mathbb{R}^n} F(y)$  alors  $F'(x; d) > 0 \quad \forall d \in \mathbb{R}^n$ .

Cette condition nécessaire affirme qu'il n'y a plus de direction de descente en  $x$ .



20

#### 4.12. Lien avec la différentiabilité.

Si  $F$  est différentiable

$$\text{Alors } F'(x; d) = \langle \nabla F(x), d \rangle \quad \forall d \in \mathbb{R}^n$$

$$\text{et } \nabla F(x) = \{ \nabla F(x) \}$$

#### 4.13. Propriétés de la fonction : $d \mapsto F'(x; d)$ .

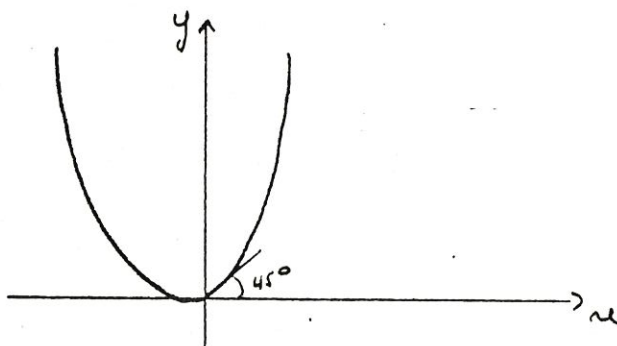
Cette fonction est convexe (donc continue) et positivement homogène.

#### Propriétés de la fonction : $x \mapsto F'(x; d)$ .

#### 4.14. Cette fonction de $x$ n'est pas localement lipschitz. contre-exemple.

Soit  $f : \mathbb{R} \rightarrow \mathbb{R}$

$$x \mapsto \begin{cases} x^2 & \text{si } x \leq 0 \\ x^2 + x & \text{si } x \geq 0. \end{cases}$$



Choisissons  $x_0 = 0$  et  $d = 1$ .

Montrons que  $f'(\cdot; d)$  n'est pas localement lipschitz en  $x_0$ .

i.e. pour tout voisinage de  $x_0$ ,  $] -r, +r[$ , avec  $r > 0$ , montrons qu'il n'existe pas de constante  $K > 0$  vérifiant :

$$|f'(x; d) - f'(y; d)| \leq K \|x - y\| \quad \text{pour tout } x \text{ et } y \in ] -r, +r[$$

Pour l'absurde, supposons qu'il existe  $K > 0$  tel que

$$|f'(x; d) - f'(y; d)| \leq K \|x - y\| \quad \text{pour tout } x \text{ et } y \in ] -r, +r[$$

Prendons  $x = -\frac{r}{m}$  et  $y = \frac{r}{m}$  où  $m \in \mathbb{N}^*$ .

D'où, pour tout  $m$ ,  $x$  et  $y \in ] -r, +r[$  et

$$\begin{aligned} |f'(x; d) - f'(y; d)| &= \left| \langle \nabla x^2 \rangle \Big|_{(-\frac{r}{m})}, 1 \rangle - \langle \nabla x^2 \rangle \Big|_{(\frac{r}{m})}, 1 \rangle \right| \\ &= \left| -\frac{2r}{m} - \frac{2r}{m} - 1 \right| \leq K \left\| \frac{2r}{m} \right\|. \end{aligned}$$

et donc  $K \geq \frac{4r + m}{2r}$  pour tout  $m$ .

Comme  $K$  est dépendant de  $m$  (donc de  $x$  et de  $y$  pris dans le voisinage de  $x_0$ ), on ne peut l'accepter.

La fonction  $x \mapsto f'(x; d)$  n'est donc pas l.l.

4.15. La fonction  $x \mapsto f'(x; d)$  n'est pas continue.

Contre exemple

Soit  $f: \mathbb{R} \rightarrow \mathbb{R}$ ,  $x \mapsto |x|$ .

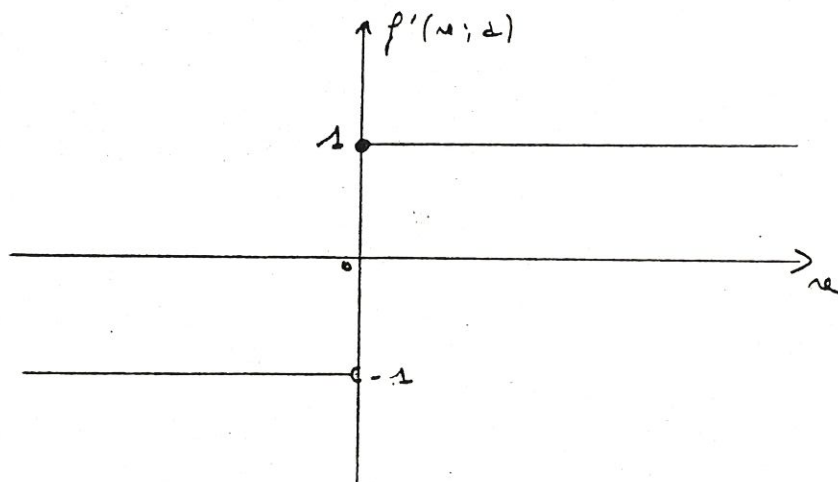
Prendons  $x_0 = 0$ .

On a  $\nabla f(x_0) = [-1, 1]$  et  $f'(x_0; d) = |d|$ .

Si  $x < 0$ , alors  $\nabla f(x) = \{ \nabla f(x) \} = \{-1\}$  et  $f'(x; d) = -d$ .

Si  $x > 0$ , alors  $\nabla f(x) = \{\nabla f(x)\} = \{1\}$  et  $f'(x; d) = d$

Fixons  $d = 1$  et représentons  $f'(\cdot; 1)$



Cette fonction est discontinue en 0. Elle ne sera donc certainement pas convexe en  $x$ .

4.16. La fonction  $x \mapsto F'(x; d)$  est semi-continue supérieurement  
([7], p. 375)

5. Sous-différentiabilité des fonctions localement lipschitzi.

5.1. Définition de la dérivée directionnelle généralisée.

La dérivée directionnelle généralisée de la fonction  $f$  en  $x_0$ , dans une direction  $\delta$  est notée  $F^\circ(x_0; \delta)$  et

$$F^\circ(x_0; \delta) = \limsup_{\substack{d \downarrow 0 \\ h \rightarrow 0}} \frac{F(x_0 + h + d\delta) - F(x_0 + h)}{d}$$

## Propositions.

Considérons une fonction  $F: \mathbb{R}^n \rightarrow \mathbb{R}$ , L.L. Dans ce cas,

5.2.  $F^\circ(x; \delta)$  existe, finie pour tout  $x$  et  $\delta \in \mathbb{R}^n$   
(contenuement à  $F'(x; \delta)$ )

5.3.  $F^\circ(x; \cdot)$  est convexe et positivement homogène en  $\delta$ .

5.4.. Si  $F'(x; \delta)$  existe,

$$\text{Alors } F'(x; \delta) \leq F^\circ(x; \delta)$$

• Si  $F$  est convexe ou différentiable

$$\text{Alors } F'(x; \delta) = F^\circ(x; \delta).$$

5.5. Définition du gradient généralisé au sens de Clarke.

Soit  $F: \mathbb{R}^n \rightarrow \mathbb{R}$  L.L. et  $x \in \mathbb{R}^n$ .

Le gradient généralisé de  $F$  en  $x$  est l'ensemble des vecteurs  $x^*$  de  $\mathbb{R}^n$  tels que

$$F^\circ(x; \delta) \geq \langle \delta, x^* \rangle \quad \text{pour tout } \delta \in \mathbb{R}^n.$$

Il est noté  $\partial^* F(x)$ .

## Propositions:

5.6.  $\partial^* F(x) = \{ x^* \mid \exists x_k \rightarrow x \text{ dans } E_1 \text{ avec } \nabla F(x_k) \rightarrow x^* \}$

où  $E_1$  est un ensemble de mesure pleine dans  $\text{dom } F$  où  $F$  est différentiable.

$$5.7. \quad F^\circ(x; \delta) = \max \{ \langle x^*, \delta \rangle \mid x^* \in \partial^* F(x) \}.$$

C'est donc la fonction d'appui du gradient généralisé de  $F$  en  $x$ .

5.8. Si  $F$ , LL. de  $\mathbb{R}^n$  dans  $\mathbb{R}$  est convexe ou différentiable,

$$\text{Alors } \partial^* F(x) = \partial F(x).$$

## 6. Pseudo et quasi convexité.

### 6.1. Définition de Pseudo-convexité.

$F$  est pseudo-convexe en  $\bar{x} \in \mathbb{R}^n$  si elle est différentiable en  $\bar{x}$  et pour  $x \in \mathbb{R}^n$ ,  $\nabla F(\bar{x})(x - \bar{x}) \geq 0$  entraîne  $F(x) \geq F(\bar{x})$ .

### 6.2. Définition de Quasi-convexité.

$F$  est quasi-convexe sur un ouvert  $\Omega$  de  $\mathbb{R}^n$  si pour tout  $x_1, x_2 \in \Omega$ ,  
 $F(dx_1 + (1-d)x_2) \leq \max \{ F(x_1), F(x_2) \}$ , avec  $d \in ]0, 1[$ .

### 6.3. Proposition.

Si  $F$  est pseudo-convexe,

Alors tout minimum local est un minimum global.



## CHAPITRE II

Etude du  $\varepsilon$ -sous-différentiel  
et de sa fonction d'appui.

Nous débuterons ce chapitre par la  
 présentation analytique et graphique du  $\varepsilon$ -sous-différentiel.  
 Deux caractérisations de cet ensemble seront développées au  
 deuxième paragraphe. Nous y définirons également la fonction  
 d'appui. Le troisième paragraphe est consacré à l'étude  
 des propriétés de la fonction  $q_{\varepsilon}(\cdot)$  définie sur  $\mathbb{R}_0^+$ , qui  
 généralise  $q(\cdot)$ , quand  $\varepsilon$  est strictement positif. Cette notion,  
 illustrée de deux exemples nous permet, au paragraphe sui-  
 vant d'étudier le comportement local du  $\varepsilon$ -sous-différentiel  
 et de démontrer le caractère localement lipschitz de sa fonction  
 d'appui. Pour terminer, nous proposons une méthode du  
 type  $\varepsilon$ -sous-gradient pour minimiser une fonction non  
 nécessairement différentiable.



## II. 1. Définition et exemples de $\varepsilon$ -sous-différentiel.

### A. Définition.

Dans le but d'élargir la notion de sous-différentiel d'une fonction  $f$ , convexe mais non nécessairement différentiable, en un point  $x_0$  du  $\text{dom } f$  (où  $\text{dom } f$  est l'ensemble sur lequel  $f$  est finie), nous nous plaçons au point  $(x_0, f(x_0) - \varepsilon)$ .

Pour chaque  $\varepsilon \geq 0$ , nous définissons le  $\varepsilon$ -sous-différentiel de  $f$  en  $x_0$  comme l'ensemble des vecteurs  $x^* \in \mathbb{R}^n$  vérifiant :

$$f(x) \geq f(x_0) + \langle x - x_0, x^* \rangle - \varepsilon \quad \text{pour tout } x \in \mathbb{R}^n \quad (2.1)$$

où  $\langle \cdot, \cdot \rangle$  désigne le produit scalaire défini sur  $\mathbb{R}^n \times \mathbb{R}^n$ .

L'ensemble de tels vecteurs, noté  $\partial_\varepsilon f(x_0)$  est un convexe fermé non vide de  $\mathbb{R}^n$ , qui est réduit au sous-différentiel  $\partial f(x_0)$  quand  $\varepsilon = 0$ .

### B. Illustration graphique.

La relation (2.1.1) signifie que l'épigraphé de la fonction affine  $H(x)$  passant par  $(x_0, f(x_0) - \varepsilon)$ , de pente  $x^*$ , contient l'épigraphé de  $f$ .

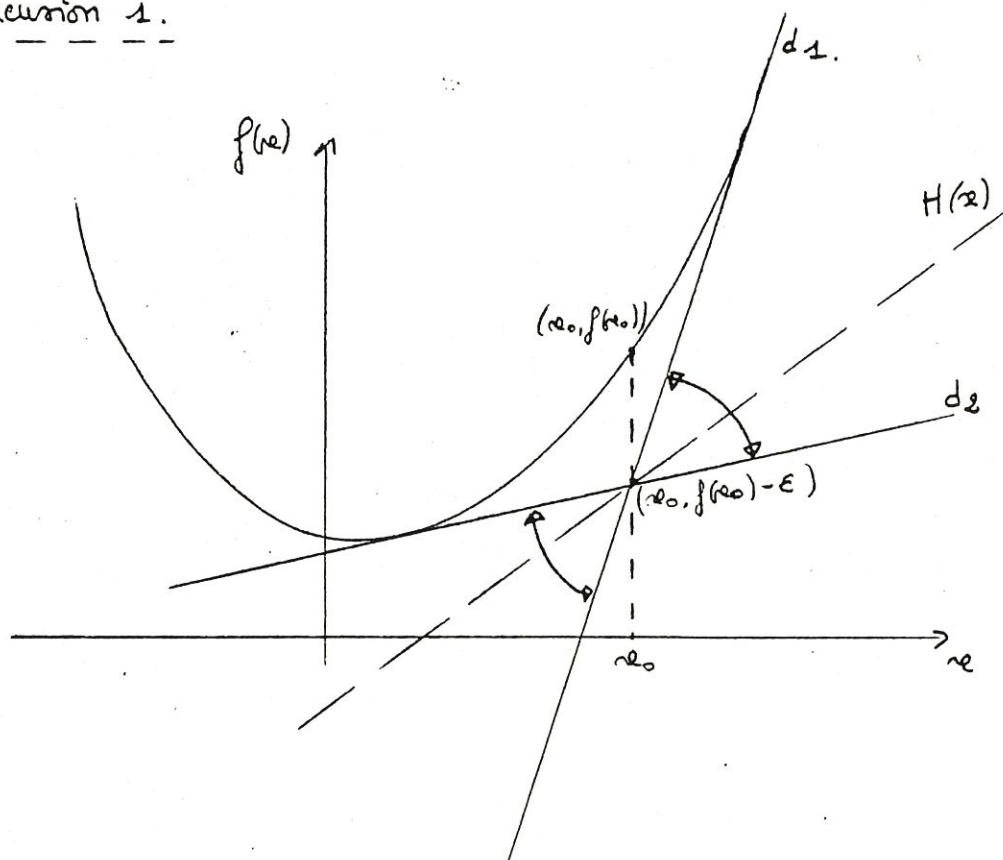
En effet, considérons  $(x, y) \in \text{épi } f$ .

$$\text{On a par définition de épi } f : f(x) \leq y$$

$$\text{Or } H(x) = \langle x^*, x - x_0 \rangle + f(x_0) - \varepsilon \leq f(x) \quad \text{par (2.1.1)}$$

d'où  $H(x) \leq y$ , ce qui signifie que  
 $(x, y) \in \text{épi } H$ , et donc  
 $\text{epi } f \subseteq \text{épi } H$ .

Ceci nous permet de construire graphiquement l'ensemble  $\partial_\varepsilon f(x_0)$   
En dimension 1.



$\partial_\varepsilon f(x_0)$  est l'ensemble de tous les coefficients angulaires compris  
entre celui de la droite  $d_1$  et celui de la droite  $d_2$ .

Remarque.

Si  $0 < \varepsilon_1 < \varepsilon_2$ , alors  $\partial f(x_0) \subseteq \partial_{\varepsilon_1} f(x_0) \subseteq \partial_{\varepsilon_2} f(x_0)$ .

Exemple 1.

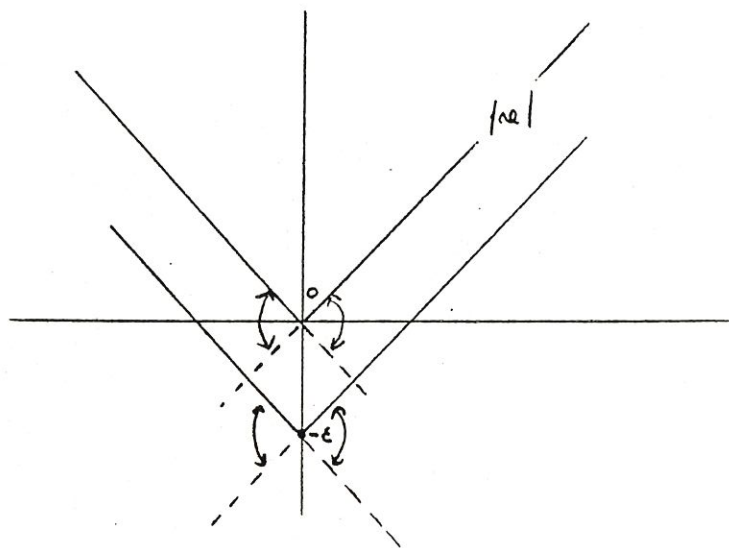
Soit  $f: \mathbb{R} \rightarrow \mathbb{R}$ ,  $x \mapsto |x| = \max(-x, x)$ .

La fonction  $f$ , comme maximum de deux fonctions affines est convexe.  
Elle n'est pas différentiable en 0.

Dans le chapitre précédent (4.2), nous avons calculé que :

$$\partial f(0) = [-1, 1].$$

Graphiquement, on peut voir ce que vaut  $\partial_\varepsilon f(0)$



Dans ce cas, on a

$$\partial_\varepsilon f(0) = \partial f(0) = [-1, 1].$$

La représentation graphique s'avère en général insuffisante pour trouver l'expression du  $\varepsilon$ -sous-différentiel, caractérisons le analytiquement.

## II.2. Caractérisation du $\varepsilon$ -sous-différentiel.

Dégageons deux formulations de  $\partial_\varepsilon f(x_0)$ . L'une, par la fonction conjuguée  $f^*$  et l'autre par la fonction d'affleurement (I,3)

## A. Caractérisation par le conjugué.

### Proposition 2.2.1

$$x^* \in \partial_\varepsilon f(x_0) \text{ si } f(x_0) + f^*(x^*) - \langle x_0, x^* \rangle \leq \varepsilon$$

Preuve.

$$\begin{aligned} x^* \in \partial_\varepsilon f(x_0) &\text{ si } \forall x \in \mathbb{R}^n : f(x) \geq f(x_0) + \langle x, x^* \rangle - \langle x_0, x^* \rangle - \varepsilon \quad (\text{déf.}) \\ &\text{ si } \forall x \in \mathbb{R}^n : \varepsilon \geq f(x_0) + \langle x, x^* \rangle - f(x) - \langle x_0, x^* \rangle \\ &\text{ si } \varepsilon \geq f(x_0) + \sup_{x \in \mathbb{R}^n} \{ \langle x, x^* \rangle - f(x) \} - \langle x_0, x^* \rangle \\ &\text{ si } \varepsilon \geq f(x_0) + f^*(x^*) - \langle x_0, x^* \rangle \quad \text{par déf. de la conj.} \end{aligned}$$

□

Remarque 1.

Comme  $f$  et  $f^*$  jouent des rôles symétriques dans la proposition 2.2.1

On a

$$x^* \in \partial_\varepsilon f(x_0) \text{ si } x_0 \in \partial_\varepsilon f^*(x^*)$$

Exemple 2.

Soit  $f$ , une fonction convexe propre et continûment homogène de  $\mathbb{R}^n$  dans  $\mathbb{R}$ .

Alors, nous montrons que  $f^*$  est la fonction indicatrice de  $\partial f(0)$ .

$$\text{i.e. } f^*(x^*) = \begin{cases} 0 & \text{si } x^* \in \partial f(0) \\ +\infty & \text{sinon.} \end{cases}$$



En effet :  $x^* \in \partial f(0)$  si  $f(x) \geq \langle x, x^* \rangle \quad \forall x \in \mathbb{R}^n$   
 si  $\langle x, x^* \rangle - f(x) \leq 0 \quad \forall x \in \mathbb{R}^n$   
 si  $\sup_{x \in \mathbb{R}^n} \{ \langle x, x^* \rangle - f(x) \} = f^*(x^*) = 0$

•  $x^* \notin \partial f(0)$  si  $\nexists x_0 \in \mathbb{R}^n$  tel que  $f(x_0) < \langle x_0, x^* \rangle$   
 si  $\nexists x_0 \in \mathbb{R}^n$  tel que  $\langle x_0, x^* \rangle - f(x_0) > 0$

$$\text{Or } \sup_{x \in \mathbb{R}^n} [\langle x, x^* \rangle - f(x)] \geq \sup_{\lambda > 0} [\langle \lambda x_0, x^* \rangle - f(\lambda x_0)]$$

et  $f$  est positivement homogène, donc  $f(\lambda x_0) = \lambda f(x_0)$ .

$$\text{dès lors } \sup_{\lambda > 0} \lambda [\langle x_0, x^* \rangle - f(x_0)] = +\infty.$$

$f^*$  est donc la fonction indicatrice de  $\partial f(0)$ , et dès lors :

$$\partial_\varepsilon f(x_0) = \{ x^* \in \partial f(0) \mid \langle x_0, x^* \rangle \geq f(x_0) - \varepsilon \}$$

### B. Caractérisation par la fonction d'affui.

Proposition 2.2.2.

Soit  $f$  une fonction convexe sur  $\mathbb{R}^n$ .

La fonction d'affui de  $\partial_\varepsilon f(x_0)$  est donnée par

$$f'_\varepsilon(x_0; \cdot) : d \longmapsto f'_\varepsilon(x_0; d) = \inf_{\lambda > 0} \frac{f(x_0 + \lambda d) - f(x_0) + \varepsilon}{\lambda}$$

Démonstration.

Au cours de cette démonstration, nous utiliserons les définitions suivantes

1) Si  $f$  est convexe, pour tout  $\lambda > 0$ , on définit :

$$(f\lambda)(x) = \lambda f(\lambda^{-1}x) \quad \text{pour tout } x \in \mathbb{R}^n.$$

2) On appelle la fonction convexe positivement homogène générée par la fonction convexe  $h$ , celle définie par :

$$f(x) = \inf \{ (h\lambda)(x) \mid \lambda > 0 \}.$$

Nous faisons également référence aux théorèmes suivants ([7], p. 118, 7)

Théorème 1.

Soit  $f$  une fonction convexe propre.

La fonction d'appui de  $\{x \mid f(x) \leq 0\}$  est  $\text{cl } g$  (I.2.7)

où  $g$  est la fonction positivement homogène générée par  $f^*$ .

De manière duale, la fermeture de la fonction positivement homogène générée par  $f$  est la fonction d'appui de  $\{x^* \mid f(x^*) \leq 0\}$

Théorème 2.

Soit  $f$ , une fonction convexe propre vérifiant  $f(0) > 0$ .

et  $h$  la fonction positivement homogène convexe générée par  $f$

Alors  $h$  est propre et  $(\text{cl } h)(x) = \inf_{\lambda > 0} (f\lambda)(x).$

Preuve de la proposition.

Si nous posons  $h(y) = f(x+y) - f(x)$ ,

nous avons :

$$h^*(x^*) = \sup_{y \in \mathbb{R}^n} \{ \langle y, x^* \rangle - h(y) \}$$

$$\begin{aligned}
 h^*(x^*) &= \sup_{y \in \mathbb{R}^n} [\langle x+y, x^* \rangle - f(x+y) - \langle x, x^* \rangle + f(x)] \\
 &= \sup_{y \in \mathbb{R}^n} [\langle x+y, x^* \rangle - f(x+y)] + f(x) - \langle x, x^* \rangle \\
 &= f^*(x^*) + f(x) - \langle x, x^* \rangle.
 \end{aligned}$$

Par la proposition 2.2.1, on a :

$$\begin{aligned}
 \cap_{\varepsilon} f(x) &= \{ x^* \mid h^*(x^*) \leq \varepsilon \} \\
 &= \{ x^* \mid h^*(x^*) - \varepsilon \leq 0 \} \\
 &= \{ x^* \mid (h+\varepsilon)^*(x^*) \leq 0 \} \quad \text{car } (h+\varepsilon)^* = h^* - \varepsilon.
 \end{aligned}$$

Par le théorème 1, on a que la fonction d'appui  $\delta^*(\cdot \mid \cap_{\varepsilon} f(x))$  est la fermeture de la fonction positivement homogène générée par  $(h+\varepsilon)^*$ .

De plus,  $(h+\varepsilon)(0) = f(x) - f(x) + \varepsilon > 0$ .

et  $(h+\varepsilon)$  est une fonction convexe propre.

On peut dès lors tirer du théorème 2 que

$$\delta^*(\cdot \mid \cap_{\varepsilon} f(x)) = \sup_{\lambda > 0} ((h+\varepsilon)\lambda)(y).$$

$$\begin{aligned}
 \text{Or } ((h+\varepsilon)\lambda)(y) &= \lambda \cdot (h+\varepsilon)(\lambda^{-1}y) \\
 &= \lambda [f(x + \lambda^{-1}y) - f(x) + \varepsilon]
 \end{aligned}$$

En échangeant  $\lambda$  et  $\lambda^{-1}$ , on obtient :

$$\delta^*(y \mid \cap_{\varepsilon} f(x)) = \sup_{\lambda > 0} \frac{f(x + \lambda y) - f(x) + \varepsilon}{\lambda} = f'_\varepsilon(x; y).$$

Et donc, la fonction d'appui du  $\varepsilon$  sous-différentiel est précisément la dérivée directionnelle à  $\varepsilon$  près.

□

Illustrons le calcul de  $f'_\varepsilon(x_0; d)$  dans deux exemples.

### ① Cas d'une fonction affine

Reprenons l'exemple 1, où  $f(x) = |x|$  et  $x_0 = 0$ .

Par la proposition 2.2.2. nous savons que

$$f'_\varepsilon(x_0; d) = \inf_{d > 0} \frac{f(x_0 + d) - f(x_0) + \varepsilon}{d}$$

Dès lors, 
$$f'_\varepsilon(0; d) = \inf_{d > 0} \frac{|d| + \varepsilon}{d} = \inf_{d > 0} |d| + \frac{\varepsilon}{d} = |d|$$

Elle nous renseigne aussi que

$$f'_\varepsilon(0; d) = \sup_{x^* \in \partial f(0) = [-1, 1]} \langle x^*, d \rangle = \begin{cases} d & \text{si } d > 0 \\ -d & \text{si } d < 0 \end{cases}$$

Un moyen de calculer  $f'_\varepsilon(x_0; d)$  est donc de chercher la fonction d'appui de  $\partial_\varepsilon f(x_0)$ .

### ② Cas d'une fonction quadratique sur $\mathbb{R}^n$ .

Exemple 3.

Définissons cette fonction comme suit:

$$f(x) = \frac{1}{2} \langle Ax, x \rangle + \langle b, x \rangle + c$$

où  $A$  est une matrice  $n \times n$ , symétrique, définie positive  
 $b$  est un vecteur de  $\mathbb{R}^n$ .

$c$  est un réel.

On calcule que

$$f'_\varepsilon(x_0; d) = \langle Ax_0, d \rangle + \langle b, d \rangle + (\varepsilon \langle Ad, d \rangle)^{1/2}.$$



En effet :

$$\begin{aligned} f'_\varepsilon(x_0; d) &= \inf_{d > 0} \frac{1}{2} \left\{ \frac{2d \langle Ax_0, d \rangle + d^2 \langle Ad, d \rangle + d \langle b, d \rangle + \varepsilon}{d} \right\} \\ &= \inf_{d > 0} \left[ \langle Ax_0, d \rangle + \frac{d}{2} \langle Ad, d \rangle + \langle b, d \rangle + \frac{\varepsilon}{d} \right] \end{aligned}$$

Annulons le dérivé du crochet pour obtenir :

$$d = \sqrt{\frac{2\varepsilon}{\langle Ad, d \rangle}}$$

et on trouve ainsi l'expression annoncée de  $f'_\varepsilon(x_0; d)$ .

De plus, la fonction  $d \mapsto (2\varepsilon \langle Ad, d \rangle)^{1/2}$  est la fonction d'appui de l'ensemble  $\{x^* \mid \langle x^*, A^{-1}x^* \rangle \leq 2\varepsilon\}$  ([7], p.1).

On peut donc écrire le  $\varepsilon$ -sous-différentiel de  $f$  en  $x_0$  comme

$$\partial_\varepsilon f(x_0) = Ax_0 + b + \{x^* \mid \langle x^*, A^{-1}x^* \rangle \leq 2\varepsilon\}$$

Comme  $f$  est différentiable,  $f'(x_0; d)$  s'exprime comme  $\langle \nabla f(x_0), d \rangle$  et donc  $f'(x_0; d) = \langle Ax_0, d \rangle + \langle b, d \rangle$ .

Remarquons que même si  $f'(x_0; d) = \lim_{\varepsilon \rightarrow 0} f'_\varepsilon(x_0; d)$ ,

la différence  $f'_\varepsilon(x_0; d) - f'(x_0; d) = \sqrt{2\varepsilon \langle Ad, d \rangle}$  décroît vers 0 aussi lentement que  $\varepsilon^{1/2}$ .

L'expression de  $f'_\varepsilon(x_0; d)$  comme  $\inf_{d > 0} \frac{f(x_0 + d) - f(x_0) + \varepsilon}{d}$

suggère l'étude de ce quotient en fonction de  $d$ .



## II. 3. Définition et propriétés de $q_{d,\varepsilon}(d)$ .

### A. Définition.

Cette fonction est définie pour un point  $x_0$ , une direction  $d$  et  $\varepsilon$  fixés comme :

$$q_{d,\varepsilon}(\cdot) : \mathbb{R}_0^+ \rightarrow \mathbb{R}$$

$$d \mapsto q_{d,\varepsilon}(d) = \frac{f(x_0 + d d) - f(x_0) + \varepsilon}{d}$$

Soulignons dès à présent que pour tout  $\varepsilon > 0$ , la fonction  $q_{d,\varepsilon}(\cdot)$  n'est pas croissante en  $d$ , contrairement à la fonction  $q_d(\cdot)$ . (I.4.9). On peut le constater aisément sur l'exemple de la fonction identité de  $\mathbb{R}$  dans  $\mathbb{R}$ .

On calcule

$$q_{d,\varepsilon}(d) = \frac{(x_0 + d d) - x_0 + \varepsilon}{d} = d + \frac{\varepsilon}{d}$$

et cette fonction n'est pas croissante en  $d$ .

### B. Comportement aux limites de $q_{d,\varepsilon}(\cdot)$ pour tout $\varepsilon \geq 0$ .

•  $\varepsilon = 0$ .

Comme la fonction  $q_d(\cdot)$  est croissante, nous avons :

$$\begin{aligned} \lim_{d \rightarrow 0^+} q_d(d) &= \sup_{d > 0} q_d(d) \\ &= f'(x_0; d). \end{aligned}$$

et 
$$\lim_{d \rightarrow +\infty} q_d(d) = \sup_{d > 0} \frac{f(x_0 + d) - f(x_0)}{d}$$

Cette limite, notée  $f_\infty(d)$ , n'est autre que la fonction de récession de  $f$ . (I.1.7)

$\varepsilon > 0$

---

$$\lim_{d \rightarrow 0^+} q_{d,\varepsilon}(d) = \lim_{d \rightarrow 0^+} \left[ q_d(d) + \frac{\varepsilon}{d} \right] = +\infty.$$

$$\lim_{d \rightarrow +\infty} q_{d,\varepsilon}(d) = \lim_{d \rightarrow +\infty} \left[ q_d(d) + \frac{\varepsilon}{d} \right] = f_\infty(d).$$

C. Propriétés de  $q_{d,\varepsilon}(\cdot)$  et définition de  $r_{d,\varepsilon}(\cdot)$ .

Comme  $f$  est convexe, la fonction  $d \mapsto f(x_0 + d)$  qui apparaît au numérateur de  $q_{d,\varepsilon}(d)$ , est également convexe.

Ceci est l'argument de preuve [développé en Annexe 1] aux assertions suivantes.

1)  $q_{d,\varepsilon}(\cdot)$  est s.e.i

2)  $q_{d,\varepsilon}(\cdot)$  est quasi-convexe sur  $\mathbb{R}_0^+$ .

3)  $q_{d,\varepsilon}(\cdot)$  est localement linéarité.

Pour caractériser un minimum de la fonction non convexe  $q_{d,\varepsilon}(\cdot)$ , nous définissons :

$$r_{d,\varepsilon}(\cdot) : \mathbb{R}_0^+ \rightarrow \mathbb{R}, \mu \mapsto q_{d,\varepsilon}(1/\mu) = r_{d,\varepsilon}(\mu).$$

Proposition 2.3.1.

La fonction  $r_{d,\varepsilon}(\cdot)$  définie sur  $\mathbb{R}_0^+$  est convexe.

Preuve.

Nous savons d'une part que :

$\sigma(\mu)$  est convexe si  $\mu \sigma(\frac{1}{\mu})$  est convexe.

et d'autre part que la fonction  $\lambda \mapsto f(x_0 + \lambda d)$  est convexe.

Donc, la fonction :  $\mu \mapsto \mu f(x_0 + \frac{1}{\mu} d)$  est convexe.

$$\begin{aligned} \text{Or } r_{d,\varepsilon}(\mu) &= \mu \left[ f(x_0 + \frac{1}{\mu} d) - f(x_0) + \varepsilon \right] \\ &= \mu f(x_0 + \frac{1}{\mu} d) + \mu [\varepsilon - f(x_0)] \end{aligned}$$

Dès lors,  $r_{d,\varepsilon}(\mu)$ , comme somme de deux fonctions convexes en  $\mu$  est aussi convexe  $\mu$ .

□

Notons que :

- Un minimum que  $q_{d,\varepsilon}(\cdot)$  sur  $\mathbb{R}_0^+$ , noté  $\lambda_0$ , équivaut au fait que  $\mu_0 = 1/\lambda_0$  est un minimum de  $r_{d,\varepsilon}(\cdot)$  sur  $\mathbb{R}_0^+$ .
- Par la caractérisation d'un minimum d'une fonction convexe, nous avons :

$$\mu_0 > 0 \text{ minimise } r_{d,\varepsilon}(\cdot) \text{ si } 0 \in \partial r_{d,\varepsilon}(\mu_0)$$

- Par une chaîne de calculs sur les gradients généralisés, nous obtenons l'égalité suivante :

$$\partial \pi_{d,\varepsilon}(\mu_0) = -\partial q_{d,\varepsilon}(d_0) \cdot d_0^2$$

Il découle alors :

$$0 \in \pi_{d,\varepsilon}(\mu_0) \text{ si } 0 \in \partial q_{d,\varepsilon}(d_0)$$

Nous pouvons donc déduire :

$$d_0 \in \mathbb{R}_0^+ \text{ est un minimum de } q_{d,\varepsilon}(\cdot) \text{ si } 0 \in \partial q_{d,\varepsilon}(d_0)$$

D. Comportements à la limite de  $\pi_{d,\varepsilon}(\mu)$ .

$$\lim_{\mu \rightarrow 0^+} \pi_{d,\varepsilon}(\mu) = \lim_{d \rightarrow +\infty} q_{d,\varepsilon}(d) = f_\infty(d)$$

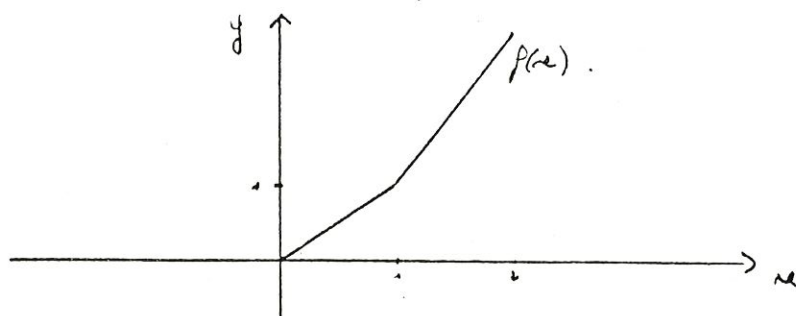
$$\lim_{\mu \rightarrow +\infty} \pi_{d,\varepsilon}(\mu) = \lim_{d \rightarrow 0} q_{d,\varepsilon}(d) = +\infty.$$

E. Exemples de calculs des fonctions  $q_{d,\varepsilon}(d)$  et  $\pi_{d,\varepsilon}(\mu)$

Exemple 4. :  $f$  est polyédrale.

Soit  $f: \mathbb{R} \rightarrow (-\infty, +\infty]$  définie par :

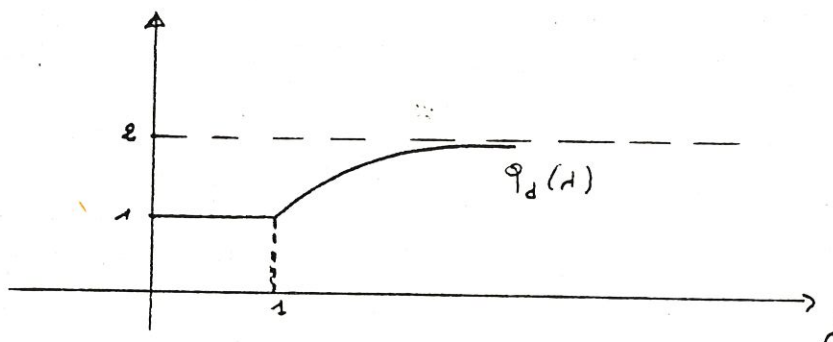
$$f(x) = \begin{cases} x & \text{si } 0 \leq x \leq 1. \\ 2x - 1 & \text{si } x \geq 1. \\ +\infty & \text{si } x < 0. \end{cases}$$



Choisissons  $x_0 = 0$  et  $d = 1$ .

Calculons d'abord :

$$q_d(\lambda) = \frac{f(x_0 + \lambda d) - f(x_0)}{\lambda} = \begin{cases} 1 & \text{si } 0 \leq \lambda \leq 1 \\ 2 - 1/\lambda & \text{si } \lambda \geq 1 \end{cases}$$



Nous vérifions que

1.  $q_d(\cdot)$  est bien croissante.
2.  $\partial f(0) = [0, 1]$  et  $f'(0; 1) = \sup_{x^* \in [0, 1]} \langle x^*, 1 \rangle = 1$ .
3.  $\lim_{\lambda \rightarrow 0^+} q_d(\lambda) = f'(0; 1) = 1$

$$\text{et } \lim_{\lambda \rightarrow \infty} q_d(\lambda) = f_\infty(1) = 2.$$

a) Fixons  $\varepsilon$  à  $1/2$

Calculons  $q_{d, \varepsilon}(\lambda)$  pour ce  $\varepsilon$

$$q_{d, \varepsilon}(\lambda) = \frac{f(\lambda) - f(0) + 1/2}{\lambda} = \begin{cases} 1 + 1/2\lambda & \text{si } 0 \leq \lambda \leq 1 \\ 2 - 1/2\lambda & \text{si } \lambda \geq 1 \end{cases}$$

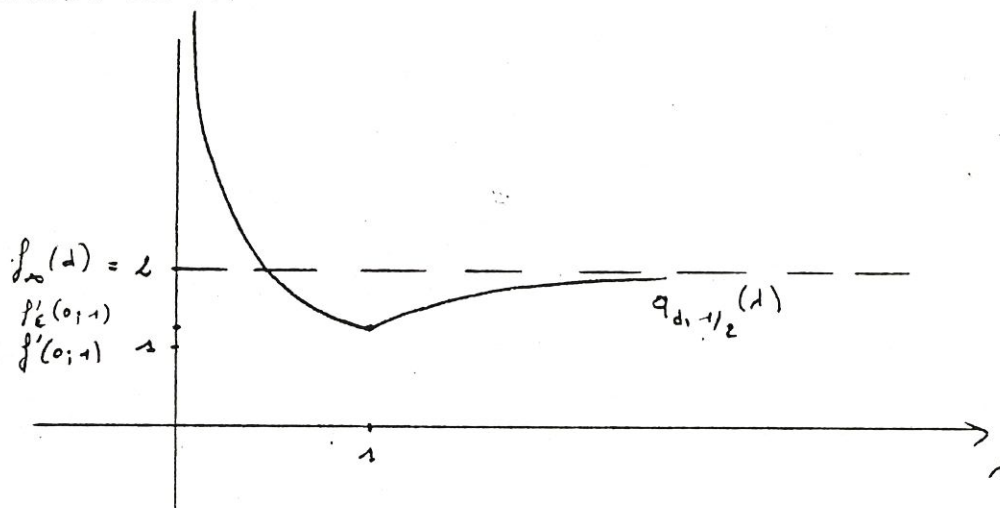
Nous vérifions que

- $\lim_{\lambda \rightarrow 0^+} q_{d, \varepsilon}(\lambda) = \lim_{\lambda \rightarrow 0^+} 1 + 1/2\lambda = +\infty$
- $\lim_{\lambda \rightarrow +\infty} q_{d, \varepsilon}(\lambda) = f_\infty(1) = 2$



$$\cdot f'_\varepsilon(0; 1) = \inf_{\lambda > 0} q_{\lambda, \varepsilon}(\lambda) = 3/2.$$

Pour se convaincre de la dernière égalité, représentons  $q_{\lambda, \varepsilon}$  en fonction de  $\lambda$ .



Remarquons que  $q_{\lambda, \varepsilon}$  n'est pas croissante, ni concave contrairement à la fonction

$$\pi_{\lambda, \varepsilon}(\mu) = q_{\lambda, \varepsilon}(1/\mu)$$

En effet, calculons

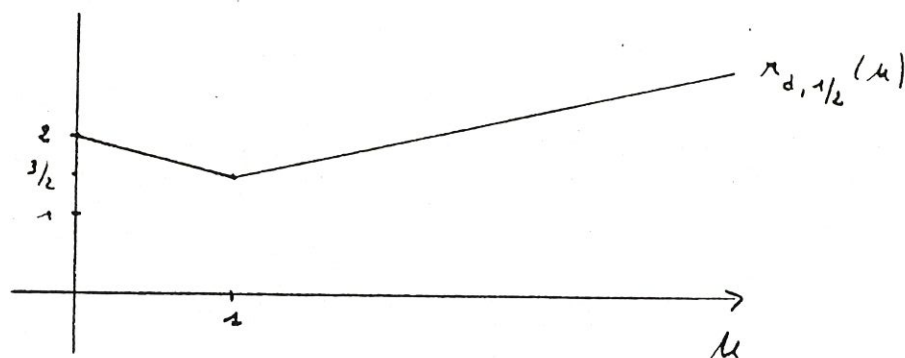
$$\pi_{\lambda, \varepsilon}(\mu) = \begin{cases} 1 + \frac{1}{2}\mu & \text{si } \mu \geq 1 \\ 2 - \frac{1}{2}\mu & \text{si } 0 \leq \mu \leq 1 \end{cases}$$

Nous constatons que

$$\cdot \pi_{\lambda, \varepsilon}(0) = f_\infty(\lambda) = 2$$

$$\cdot \lim_{\mu \rightarrow +\infty} \pi_{\lambda, \varepsilon}(\mu) = \lim_{\mu \rightarrow +\infty} 1 + \frac{1}{2}\mu = +\infty$$

Le graphe de la fonction  $\pi_{\lambda, \varepsilon}(\mu)$  confirme la convexité.



Remarquons que l'élément  $\mu_0 (=1)$  qui minimise  $r_{d, \varepsilon}$  est bien  $1/\mu_0$  si  $\mu_0$  est évalué précédemment à 1.

b) Pour  $\varepsilon$  arbitraire

Calculons

$$q_{d, \varepsilon}(h) = \frac{f(\mu_0 + h) - f(\mu_0) + \varepsilon}{h} = \begin{cases} 1 + \varepsilon/h & \text{si } 0 < h \leq 1 \\ 2 + (\varepsilon - 1)/h & \text{si } h \geq 1 \end{cases}$$

Distinguons trois cas suivant la valeur de  $\varepsilon$ .

1)  $\varepsilon < 1$ .

Ce cas est illustré par l'exemple où  $\varepsilon = 1/2$ .

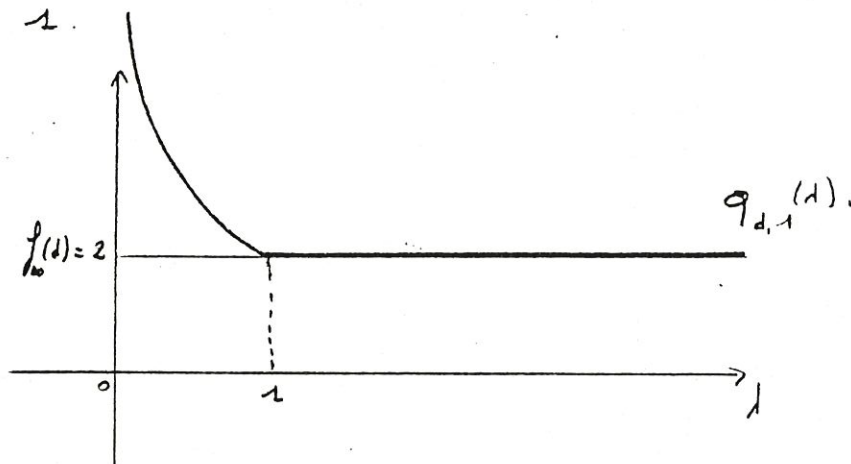
Nous observons que la fonction  $q_{d, \varepsilon}(h)$  est décroissante sur  $[0, 1]$  et croissante sur  $[1, +\infty[$ . Le minimum est donc atteint par  $h = 1$  et vaut  $f'_\varepsilon(0; 1) = \varepsilon + 1$ .

Soulignons que  $\lim_{\varepsilon \rightarrow 0^+} f'_\varepsilon(0; 1) = f'(0; 1) = 1$ .

2)  $\varepsilon = 1$ .

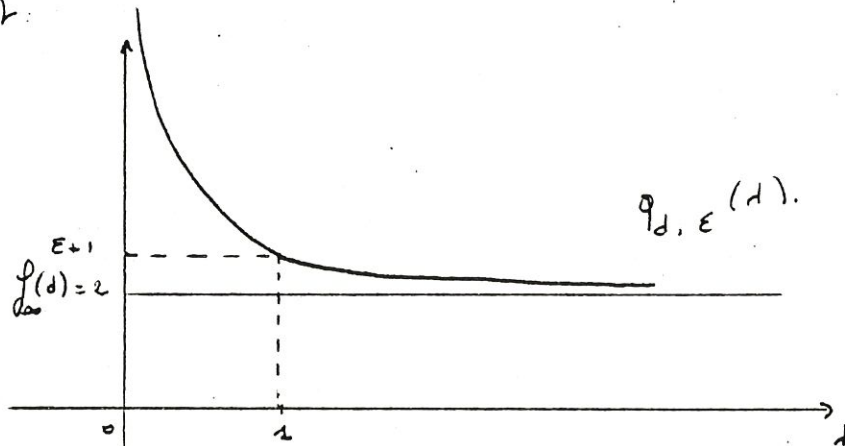
$$q_{d, \varepsilon}(h) = \begin{cases} 1 + 1/h & \text{si } 0 < h \leq 1 \\ 2 & \text{si } h \geq 1. \end{cases}$$

Si l'on représente cette fonction, on s'aperçoit que son infimum  $f'_\varepsilon(\infty; d)$  aussi égal à  $f_\infty(d)$  est atteint pour tout  $d \geq 1$ .



3)  $\varepsilon > 1$ .

La fonction  $q_{d,\varepsilon}(d)$  est décroissante sur  $\mathbb{R}_0^+$  et se représente comme suit:



Elle reste supérieure à la droite  $f_\infty(d)$  pour tout  $d > 0$ . Son minimum  $f'_\varepsilon(\infty; d) = f_\infty(d)$  n'est pas atteint sur  $\mathbb{R}_0^+$ .

En terme de fonction  $r_{d,\varepsilon}$ , cela signifie que  $\mu_0 = 0$  est l'unique minimum de  $r_{d,\varepsilon}$  sur  $\mathbb{R}^+$ . Il est donc naturel de poser  $r_{d,\varepsilon}(0) = f_\infty(d)$  chaque fois que  $f_\infty(d) < +\infty$ .

### Conclusions de cet exemple.

Dans les trois situations précédentes, l'ensemble des  $d$  pour lesquels  $q_{d,\varepsilon}(d)$  atteint son infimum  $f'_\varepsilon(\alpha_0; d)$ , présente des caractéristiques différentes, à savoir, être borné, non borné ou vide. Cette distinction sera capitale pour la suite. Il est donc utile de définir rigoureusement:

1) L'ensemble des  $d_0 \in \mathbb{R}_0^+$  qui minimisent  $q_{d,\varepsilon}(d)$ :

$$\Lambda_{d,\varepsilon}(\alpha_0) = \{d_0 \in \mathbb{R}_0^+ \mid q_{d,\varepsilon}(d_0) = f'_\varepsilon(\alpha_0; d)\}$$

2) L'ensemble des  $\mu_0 \in \mathbb{R}^+$  qui minimisent  $r_{d,\varepsilon}(\mu)$ :

$$\Pi_{d,\varepsilon}(\alpha_0) = \{\mu_0 \in \mathbb{R}^+ \mid \mu_0 = 1/d_0, d_0 \in \Lambda_{d,\varepsilon}(\alpha_0)\}$$

Cet exemple nous a permis de constater que  $\Lambda_{d,\varepsilon}(\alpha_0)$  peut être vide et dans ce cas,  $\Pi_{d,\varepsilon}(\alpha_0)$  se réduit au singleton  $\{0\}$ . Dans le cas contraire,  $\Lambda_{d,\varepsilon}(\alpha_0)$  peut être un intervalle du type  $[d_0, +\infty[$ . Dès lors,  $\Pi_{d,\varepsilon}(\alpha_0)$  contient l'élément 0 mais n'y est pas réduit,  $[0, 1/d_0]$  (avec comme convention:  $1/\infty = 0$  et  $1/0 = +\infty$ ).

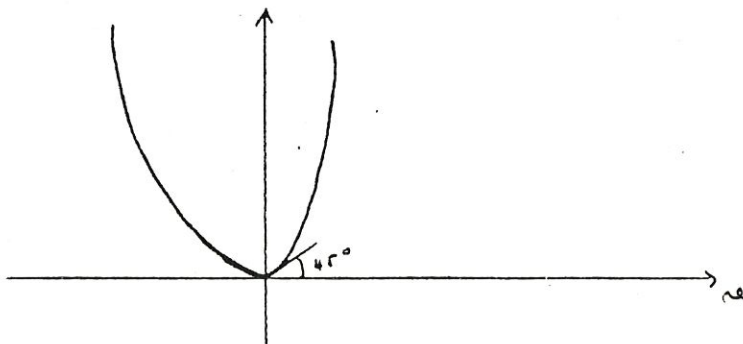
Remarquons que  $\Pi_{d,\varepsilon}(\alpha_0)$  ainsi défini est un intervalle compact, non vide de  $\mathbb{R}^+$ , pour tout  $\varepsilon > 0$ .

### Remarques.

1. A la lumière de l'exemple 4, une condition nécessaire et suffisante pour que  $q_{d,\varepsilon}$  atteigne son minimum sur  $\mathbb{R}_0^+$  est de supposer qu'il existe  $d_* > 0$  tel que  $q_{d,\varepsilon}(d_*) \leq f_\infty(d)$ . Cette condition est automatiquement satisfaite pour les directions  $d$  telles que  $f_\infty(d) = +\infty$ .
2. Par la définition même de l'ensemble  $\Pi_{d,\varepsilon}(x)$  et les propriétés de continuité de  $f$ , la multiaffiliation  $x \mapsto \Pi_{d,\varepsilon}(x)$  est semi-continue supérieurement et bornée dans un voisinage de  $x_0$ .
3. Quand  $\varepsilon = 0$ , les ensembles  $\Lambda_{d,\varepsilon}(x)$  et  $\Pi_{d,\varepsilon}(x)$  sont notés respectivement  $\Lambda_d(x)$  et  $\Pi_d(x)$ .

### Exemple 5.

Considérons une fonction quadratique par morceaux de  $\mathbb{R}$  dans  $\mathbb{R}$  telle que  $f(x) = \begin{cases} x^2 & \text{si } x \leq 0 \\ x^2 + x & \text{si } x \geq 0 \end{cases}$ .



Choisissons  $x_0 = 1$  et  $d = 1$ .



Calculons

$$q_d(\lambda) = \frac{f(\lambda) - f(0)}{\lambda} = \lambda + 1$$

$$f_\infty(d) = \lim_{\lambda \rightarrow +\infty} q_d(\lambda) = +\infty$$

$$f'(0; 1) = \inf_{\lambda > 0} q_d(\lambda) = 1.$$

Il n'y a pas de  $\lambda > 0$  qui minimise  $q_d(\cdot)$ .  $\Lambda_d(\cdot)$  est donc vide.

Nous calculons ensuite que

$$q_{d,\varepsilon}(\lambda) = \lambda + 1 + \frac{\varepsilon}{\lambda}$$

Calculons la dérivée de cette fonction pour trouver  $\Lambda_{d,\varepsilon}(x_0)$  et  $f'_\varepsilon(0; 1)$ .

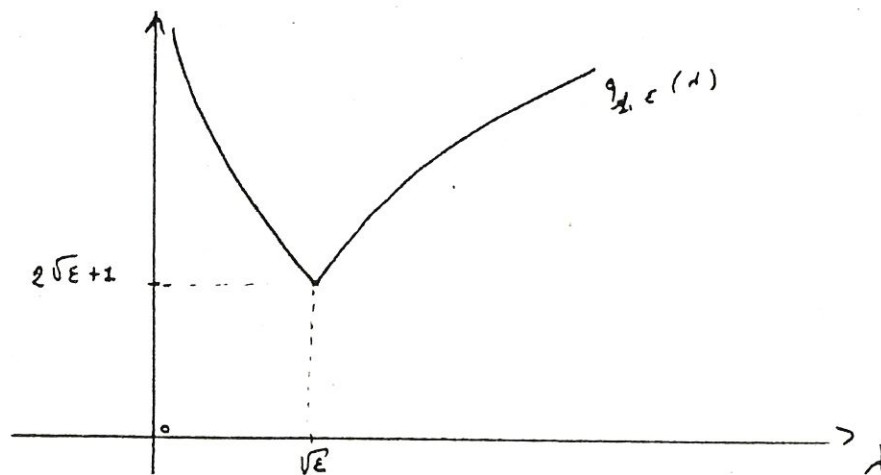
$$(q_{d,\varepsilon}(\lambda))' = 1 - \frac{\varepsilon}{\lambda^2} = 0 \text{ si } \lambda = \sqrt{\varepsilon}.$$

Dès lors  $\Lambda_{d,\varepsilon}(0) = \{\sqrt{\varepsilon}\}$ , et

$$f'_\varepsilon(0; 1) = 2\sqrt{\varepsilon} + 1$$

Comme annoncé dans la remarque 1,  $\Lambda_{d,\varepsilon}(0)$  est non vide puisque  $f_\infty(d) = +\infty$ , et donc

$$q_{d,\varepsilon}(\lambda) \leq f_\infty(d).$$



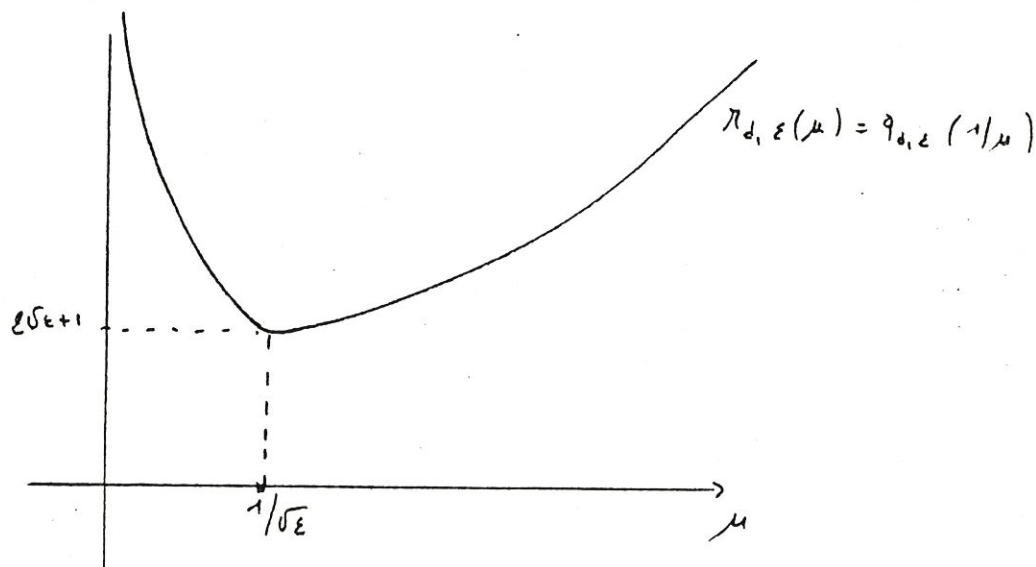
Examinons maintenant le comportement de la fonction  $r_{d,\varepsilon}(\cdot)$  et l'ensemble  $\Pi_{d,\varepsilon}$ .

$$r_{d,\varepsilon}(\mu) = \frac{1}{\mu} + 1 + \varepsilon\mu.$$

$$\lim_{\mu \rightarrow \infty} r_{d,\varepsilon}(\mu) = +\infty$$

Le minimum de  $r_{d,\varepsilon}(\mu)$  est atteint par  $\mu = 1/\sqrt{\varepsilon}$  et vaut  $2\sqrt{\varepsilon} + 1 (= f'_\varepsilon(0; 1))$ .

$$\text{D'où, } \Pi_{d,\varepsilon}(0) = 1/\Lambda_{d,\varepsilon}(0) = \left\{ \frac{1}{\sqrt{\varepsilon}} \right\}.$$



Signalons que des règles de calculs sur  $f'_\varepsilon(\alpha_0; d)$  tels que la dérivée directionnelle d'une somme, d'une composée, d'un maximum de fonctions convexes sont énoncées en annexe 2.

Les fonctions  $q_{d,\varepsilon}(\cdot)$  et  $r_{d,\varepsilon}(\cdot)$  étudiées dans ce paragraphe nous permettent d'examiner le comportement de  $f'_\varepsilon(\alpha_0; d)$  comme fonction de  $\varepsilon$  et de  $\alpha$ .

## II.4. Comportement local de $f'_\varepsilon(x_0; d)$

Noté but, dans ce paragraphe, est de prouver la caractéristique localement lipschitz de  $f'_\varepsilon(x; d)$ , fonction d'appui de  $\partial_\varepsilon f(x)$ . Il est donc nécessaire d'étudier les propriétés locales de la multiplication  $\partial \cdot f(\cdot)$ .

Nous savons qu'elle est à valeur dans la collection des ensembles convexes fermés de  $\mathbb{R}^n$ , que nous notons  $\mathcal{C}(\mathbb{R}^n)$ .

La question est donc de savoir comment varie

$$\partial \cdot f(\cdot) : \mathbb{R}^n \times \mathbb{R}^+ \longrightarrow \mathcal{C}(\mathbb{R}^n)$$

$$(x, \varepsilon) \longmapsto \partial_\varepsilon f(x), \text{ aux variables } x \text{ et } \varepsilon.$$

Pour mesurer cette variation, on introduit la topologie de Hausdorff,

où pour chaque ensemble convexe fermé  $C$  de  $\mathcal{C}(\mathbb{R}^n)$ , les systèmes fondamentaux de voisinages sont donnés par les ensembles de la forme

$$\{ D \in \mathcal{C}(\mathbb{R}^n) \mid D \subset C + \alpha B^* \text{ et } C \subset D + \alpha B^* \}$$

où  $\alpha$  parcourt  $\mathbb{R}_0^+$  et  $B^*$  est la boule unité de  $\mathbb{R}^n$ .

Le résultat suivant assure la continuité du sous-différentiel approximant. ([13], p. 456-458)

### Théorème 2.1.

Soit  $f$  une fonction convexe propre de  $\mathbb{R}^n$  dans  $\mathbb{R}$ .

Alors, la multiplication  $(x, \varepsilon) \longmapsto \partial_\varepsilon f(x)$  est continue sur  $\mathbb{R}^n \times \mathbb{R}_0^+$  au sens de la topologie de Hausdorff.

Dans le théorème suivant, nous "quantifions" cette continuité à l'aide de la métrique définie par l'égalité de Hölder :

$$h(C, D) = \sup_{\|d\| \leq 1} |\delta^*(d|C) - \delta^*(d|D)|$$

$$\text{où } \delta^*(d|C) = \sup_{y \in C} \langle y, d \rangle.$$

Théorème 2.2.

Soit  $f$  convexe de  $\mathbb{R}^n$  dans  $\mathbb{R}$ .

Alors il existe une constante  $k$  telle que

$$h(\partial_\varepsilon f(x), \partial_{\varepsilon'} f(x')) \leq \frac{k}{\min(\varepsilon, \varepsilon')} (\|x - x'\| + |\varepsilon - \varepsilon'|)$$

pour tout  $x, x' \in \mathbb{R}^n$  et  $\varepsilon, \varepsilon' \in \mathbb{R}_0^+$ .

Preuve.

Soit  $d$  tel que  $\|d\| \leq 1$ .

Etant donné  $x \in \mathbb{R}^n$  et  $\varepsilon \in \mathbb{R}^+$ , on sait que la fonction  $q_{d, \varepsilon}(t) = \frac{f(x + td) - f(x) + \varepsilon}{t}$  atteint son minimum  $f'_\varepsilon(x; d)$  sur  $\mathbb{R}$ .

Comme  $f$  est convexe, elle est localement lipschitz, de constante  $\alpha$ . Il est prouvé en annexe 3 que pour tout  $\alpha$  et  $\varepsilon > 0$ , tous les  $d_\alpha$  qui satisfont :

$$f'_\varepsilon(x; d) \leq q_{d, \varepsilon}(d_\alpha) \leq f'_\varepsilon(x; d) + \alpha \quad (1)$$

appartiennent à l'intervalle  $\left[ \frac{\varepsilon}{2\alpha + \alpha}, +\infty \right]$ , quelque soient  $x \in \mathbb{R}^n$  et  $d$  tel que  $\|d\| \leq 1$ .



Calculons

$$\begin{aligned} h(\partial_\varepsilon f(x), \partial_{\varepsilon'} f(x')) &= \sup_{\|d\| \leq 1} |\mathcal{S}^*(d | \partial_\varepsilon f(x)) - \mathcal{S}^*(d | \partial_{\varepsilon'} f(x'))| \\ &= \sup_{\|d\| \leq 1} |f'_\varepsilon(x; d) - f'_{\varepsilon'}(x'; d)| \end{aligned}$$

On a toujours :  $f'_\varepsilon(x; d) \leq \varphi_{d, \varepsilon}(d)$  pour tout  $d > 0$ .

D'autre part, pour tout  $\alpha' > 0$  et  $d \geq \frac{\varepsilon'}{2r + \alpha'}$ , on a par (1)

$$-f'_{\varepsilon'}(x'; d) \leq -\varphi_{d, \varepsilon'}(d) + \alpha'.$$

Donc

$$\begin{aligned} f'_\varepsilon(x; d) - f'_{\varepsilon'}(x'; d) &\leq \frac{f(x+ld) - f(x) + \varepsilon}{d} - \frac{f(x'+ld) - f(x') + \varepsilon'}{d} + \alpha' \\ &= \frac{f(x+ld) - f(x'+ld)}{d} + \frac{f(x') - f(x)}{d} + \frac{\varepsilon - \varepsilon'}{d} + \alpha' \end{aligned}$$

ceci pour tout  $\alpha' > 0$  et  $d \geq \frac{\varepsilon'}{2r + \alpha'}$ .

Comme  $f$  est L.L. de constante  $r$ , on a :

$$\|f(x+ld) - f(x'+ld)\| \leq r \|x - x'\|$$

$$\|f(x) - f(x')\| \leq r \|x - x'\|.$$

$$\text{D'où } f'_\varepsilon(x; d) - f'_{\varepsilon'}(x'; d) \leq \frac{2r \|x - x'\| + |\varepsilon - \varepsilon'|}{d} + \alpha' \quad (2)$$

pour tout  $\alpha' > 0$  et  $d \geq \frac{\varepsilon'}{2r + \alpha'}$

Par le même raisonnement, on obtient :

$$f'_{\varepsilon'}(x'; d) - f'_\varepsilon(x; d) \leq \frac{2r \|x - x'\| + |\varepsilon - \varepsilon'|}{d} + \alpha \quad (3)$$

pour tout  $\alpha > 0$  et  $d \geq \frac{\varepsilon}{2r + \alpha}$



Faisons  $\alpha'$  et  $\alpha$  tendre vers 0 dans (2) et (3), et posons  $k = \max(4r^2, 2r)$

On obtient donc :

$$|f'_\varepsilon(x; d) - f'_{\varepsilon'}(x'; d)| \leq \frac{k}{\min(\varepsilon, \varepsilon')} (\|x - x'\| + |\varepsilon - \varepsilon'|)$$

où  $|f'_\varepsilon(x; d) - f'_{\varepsilon'}(x'; d)|$  est précisément  $h(\partial_\varepsilon f(x), \partial_{\varepsilon'} f(x'))$ , ce qui termine la démonstration.  $\square$

Écrivons maintenant les deux résultats principaux :

Corollaire 1. : Comportement de  $f'_\varepsilon(x; d)$  comme fonction de  $\varepsilon$ .

Soit  $f$  une fonction convexe de  $\mathbb{R}^n$  dans  $\mathbb{R}$ .

Soit  $x_0$  un point fixé de  $\mathbb{R}^n$ .

Alors, pour tout  $\bar{\varepsilon} > 0$ , il existe une constante  $k$  tq

$$|f'_\varepsilon(x_0; d) - f'_{\varepsilon'}(x_0; d)| \leq \frac{k}{\min(\varepsilon, \varepsilon')} |\varepsilon - \varepsilon'|$$

pour tout  $\varepsilon, \varepsilon' \in ]0, \bar{\varepsilon}[$ .

On obtient cette inégalité en posant  $x = x'$  dans le théorème 2.2.

Corollaire 2. : Comportement de  $f'_\varepsilon(x; d)$  comme fonction de  $x$ .

Soit  $x_0$  un point fixé de  $\mathbb{R}^n$ .

Alors, il existe un voisinage  $V$  de  $x_0$  et une  $c \stackrel{\text{def}}{=} \frac{1}{k} > 0$  tq

$$|f'_\varepsilon(x; d) - f'_\varepsilon(x'; d)| \leq \frac{k}{\varepsilon} \|x - x'\| \text{ pour tout } x, x' \in V$$

ie.  $f'_\varepsilon(x; d)$  est localement Lipschitz en  $x$ .

On obtient cette inégalité en posant  $\varepsilon = \varepsilon'$  dans le théorème 2.2

Notons la fonction  $x \mapsto f'_\varepsilon(x; d)$  par  $v_{d, \varepsilon}(x)$ . Celle-ci étant l.l., on peut dès lors parler de son sous-différentiel au sens de Clarke (I.5) et étudier ses points de différentiabilité. Ceci fera l'objet du chapitre III. Dans cette partie, nous développerons l'étude de la dérivée directionnelle généralisée de  $v_{d, \varepsilon}(x)$ . Grâce à cet outil du second ordre, nous espérons améliorer les algorithmes utilisant la dérivée directionnelle du premier ordre, pour minimiser une fonction non nécessairement différentiable.

Auparavant, nous présentons une méthode numérique de descente pour une fonction objective de ce type.

### III. 5. Exemple d'algorithme du premier ordre.

Habituellement, les algorithmes de descente nécessitent l'existence du gradient de la fonction de coût, car il intervient dans les calculs et dans la preuve de la convergence de la méthode vers un minimum local (et donc global).

Mais dans beaucoup de problèmes, la fonction objective est convexe et non nécessairement différentiable. C'est ce type de problèmes que nous résolvons par la méthode du  $\varepsilon$  sous-gradient.

Considérons le problème  $\begin{cases} \min g(x) \\ \text{s.t. } x \in X \end{cases}$   
 où  $g$  est convexe et  $X \subset \mathbb{R}^n$ .

Plutôt que de considérer explicitement les contraintes, on pose

$$f(x) = g(x) + \delta(x|X) \quad \text{où} \quad \delta(x|X) = \begin{cases} 0 & \text{si } x \in X \\ +\infty & \text{sinon.} \end{cases}$$

Le problème revient donc toujours à minimiser  $f$  de  $\mathbb{R}^n$  dans  $]-\infty, +\infty]$   
 avec  $f$  convexe et  $f(x) < +\infty$  pour au moins un  $x$ .

L'algorithme se base sur les trois propositions suivantes :

Proposition 2.5.1.

Pour tout  $\varepsilon_1, \varepsilon_2 \in \mathbb{R}_0^+$

Si  $0 < \varepsilon_1 < \varepsilon_2$  Alors  $\partial f(x) \subseteq \partial_{\varepsilon_1} f(x) \subseteq \partial_{\varepsilon_2} f(x)$

Proposition 2.5.2.

$$0 \in \partial_{\varepsilon} f(x) \text{ si } 0 \leq f(x) - \inf_{z \in \mathbb{R}^n} f(z) \leq \varepsilon$$

Proposition 2.5.3.

Soit  $x$  tel que  $f(x) < \infty$  et  $0 \notin \partial_{\varepsilon} f(x)$

Soit  $y$  tel que  $\sup_{x^* \in \partial_{\varepsilon} f(x)} \langle y, x^* \rangle < 0$  (2.5.1)

Alors  $f(x) - \inf_{d \geq 0} f(x + dy) > \varepsilon$ .

Comment trouver  $y$  vérifiant (2.5.1) ?

Soit  $\bar{x}^*$  l'unique vecteur de norme minimale de  $\partial_\varepsilon f(x)$   
 Alors  $y = -\frac{\bar{x}^*}{\|\bar{x}^*\|}$  satisfait (2.5.1)

En effet :  $\sup_{x \in \partial_\varepsilon f(x)} \langle -\frac{\bar{x}^*}{\|\bar{x}^*\|}, x^* \rangle = -\|\bar{x}^*\| < 0$

Algorithme  $\varepsilon$  sous-gradient.

Initialisation.

choisir  $x_0$  tq  $f(x_0) < +\infty$ .

$\varepsilon_0 > 0$

$\alpha \in \mathbb{R}$ ,  $0 < \alpha < 1$ .

Étape principale.

#1. Soient  $x_n$  et  $\varepsilon_n > 0$

Poser  $\varepsilon_{n+1} = \alpha^k \varepsilon_n$ , où  $k$  est le plus petit entier  $\geq 0$

tq  $0 \notin \partial_{\varepsilon_{n+1}} f(x_n)$ .

#2. Trouver  $y_n$  tq  $\sup_{x^* \in \partial_{\varepsilon_{n+1}} f(x_n)} \langle y_n, x^* \rangle < 0$

#3. Poser  $x_{n+1} = x_n + \lambda_n y_n$  où  $\lambda_n > 0$  est tq

$f(x_n) - f(x_{n+1}) > \varepsilon_{n+1}$

aller en #1.



### Remarques concernant l'algorithme.

1. Si  $x_m$  n'est pas un point minimum de  $f$ , il existe toujours  $k \geq 0$  tel que  $0 \notin \partial_{\alpha^k \varepsilon_m} f(x_m)$   
 en effet, par la proposition 2.5.2, on a :  
 $0 \notin \partial_{\alpha^k \varepsilon_m} f(x_m)$  si  $f(x_m) - \inf_x f(x) > \alpha^k \varepsilon_m$ .

2. La proposition 2.5.3 nous assure qu'il existe  $d_m$  tel que  
 $f(x_m) - f(x_m + d_m x_m) > \varepsilon_{m+1}$

Le pas #3 peut donc toujours être effectué. De plus, on peut montrer que l'ensemble  $\{d_m \mid f(x_m) - f(x_m + d_m x_m) > \varepsilon_{m+1}\}$  est un intervalle borné ou une demi-droite ouverte. Une manière de trouver  $d_m$  est donc d'employer la recherche linéaire :

$$f(x_m + d_m y_m) = \min_{d > 0} f(x_m + d y_m)$$

3. La convergence de cette méthode est prouvée en ([6]), où les aspects techniques sont également envisagés.

Dans ce chapitre, nous nous sommes familiarisés avec la notion de  $\varepsilon$ -sous-différentiel, ses caractérisations, avec les fonctions  $q_{d,\varepsilon}(\cdot)$  et  $\pi_{d,\varepsilon}(\cdot)$ . Nous avons esquisé l'utilité du  $\partial_\varepsilon f(x)$  dans un algorithme. Nous avons aussi mis l'accent sur le caractère localement lipschitz de la fonction d'appui de l'ensemble  $\partial_\varepsilon f(x)$ , à savoir  $f'_\varepsilon(x; d)$ , aussi notée  $V_{d,\varepsilon}(x)$ , propriété dont ne jouissait pas  $f'(x; d)$ . Ceci est un outil essentiel pour étudier la différentiabilité de  $V_{d,\varepsilon}(\cdot)$



### CHAPITRE III

Etude de la différentiabilité  
de  $V_{d,\varepsilon}(\cdot)$  .

Dans ce chapitre, nous étudierons sous quelles conditions la fonction  $v_{d,\varepsilon} : x \rightarrow f'_\varepsilon(x; d)$  est différentiable. Pour cela, nous commencerons par montrer, dans un premier paragraphe, que  $v_{d,\varepsilon}$  admet une dérivée directionnelle en chaque point  $x_0$  dans toute direction  $\delta$ . Dans le second paragraphe, nous caractériserons le gradient généralisé de la fonction localement lipschitz  $v_{d,\varepsilon}$  et nous en déduirons des conditions assurant la différentiabilité de celle-ci. Enfin, dans un troisième paragraphe, nous étudierons plus en détail cette différentiabilité lorsque l'ensemble  $\Lambda_{d,\varepsilon}(x_0)$  est respectivement non vide et borné, non borné, vide. Des propriétés particulières seront déduites et illustrées par des exemples. Dans tout ce chapitre, la direction  $d$  et le paramètre  $\varepsilon > 0$  sont fixés.

## III. 1

Expression de la dérivée directionnelle  
de  $v_{d,\varepsilon}(\cdot)$  dans une direction  $\delta$

---

Nous allons examiner ce qui se passe  
autour de  $x_0$  en calculant la dérivée directionnelle  
de la fonction  $v_{d,\varepsilon} : x_0 \rightarrow f'_\varepsilon(x_0; d)$  et, pour cela, nous  
nous référons au théorème suivant.

Théorème 3.1

Pour toute direction  $\delta$ ,  
le quotient  $\frac{f'_\varepsilon(x_0 + s\delta; d) - f'_\varepsilon(x_0; d)}{s}$   
admet une limite  $v'_{d,\varepsilon}(x_0; \delta)$  quand  $s \rightarrow 0$ ,  
aussi notée  $f''_\varepsilon(x_0; d; \delta)$  et

$$f''_\varepsilon(x_0; d, \delta) = \min_{\eta \in M_{d,\varepsilon}(x_0)} \max_{x^* \in \partial_\varepsilon f(x_0)_d} \{ \eta[x^*, \delta] - f'_\varepsilon(x_0; \delta) \}$$

(3.1.1)

où  $\partial_\varepsilon f(x_0)_d = \{ x^* \in \partial_\varepsilon f(x_0) \text{ t.q. } \langle x^*, d \rangle = f'_\varepsilon(x_0; d) \}$ .  
De plus, les opérations max et min commutent.

Avant de prouver ce théorème, introduisons quelques notations et définitions.

### Notations et définitions

- Notons par  $\partial_\varepsilon f(y)_d$  l'ensemble des solutions optimales du problème d'optimisation  $P_y$ , pour  $y$  fixé,

$$P_y : \begin{aligned} & \max (x^*, d) \\ & \text{s.t. } x^* \in \partial_\varepsilon f(y) \end{aligned}$$

et notons par  $v_{d,\varepsilon}(y)$  ce maximum.

Nous pouvons donc écrire que

$$\partial_\varepsilon f(y)_d = \{ x^* \in \partial_\varepsilon f(y) \text{ t.q. } v_{d,\varepsilon}(y) = \langle x^*, d \rangle \}$$

- De même, notons  $U(y)$  l'ensemble des solutions optimales du problème dual au problème  $P_y$ , appelé  $Q_y$ , pour  $y$  fixé.

$$Q_y : \inf_{u \geq 0} \theta(y, u).$$

$$\text{où } \theta(y, u) = \sup \{ L(x^*, y, u) \mid x^* \in \mathbb{R}^n \}$$

$$\text{et } L(x^*, y, u) = \langle d, x^* \rangle - u [f(y) + f^*(x^*) - \langle x^*, y \rangle - \varepsilon]$$

Appelons  $d(y)$  cet infimum.

Nous pouvons donc tirer que:

$$U(y) = \{ u \geq 0 \mid d(y) = \theta(y, u) \}$$

On remarque immédiatement que, puisque  $\partial_\varepsilon f(y)$  est un ensemble compact, convexe et non vide,  $\partial_\varepsilon f(y)_d$  est aussi un ensemble compact, convexe, non vide.

De plus, puisque  $\partial_\varepsilon f(\cdot)$  est localement lipschitz,

$\partial_\varepsilon f(\cdot)_d$  est fermé et localement borné.

Rappelons que  $\partial_\varepsilon f(\cdot)_d$  est localement borné en  $y$  signifie qu'il existe un voisinage  $V$  de  $y$  tel que

$$\bigcup_{z \in V} \partial_\varepsilon f(z)_d \text{ est bornée}$$

Dans la démonstration du théorème, nous ferons appel à un lemme qui caractérise l'ensemble  $U(y)$ .

Lemme 3.1

$$1) \quad d(y) = v_{d, \varepsilon}(y)$$

$U(y)$  est un ensemble convexe, compact, non vide t.q

a) chaque couple  $(\hat{x}^*, \hat{u})$  avec  $\hat{u} \in U(y), \hat{x}^* \in \partial_\varepsilon f(y)_d$  est un point selle de la fonction  $L(y, \cdot, \cdot)$  et

$$L(y, \hat{x}^*, \hat{u}) = v_{d, \varepsilon}(y)$$



b) Si  $(f^+)_d(d) = +\infty$  alors  $0 \notin U(y)$  et  
$$U(y) = \{u > 0 : u[f(y + \frac{d}{u}) - f(y) + \varepsilon] = v_{d,\varepsilon}(y)\}$$
où  $(f^+)_d$  est la fonction récession de  $f$ .  
2)  $U(\cdot)$  est une multiapplication localement bornée

Vu l'aspect technique de la démonstration de ce lemme ([3], p.3), nous préférons nous attarder sur la preuve du théorème

Preuve du théorème

Dans la suite, pour alléger l'écriture, nous noterons par  $M(\cdot)$  la multiapplication  $\partial_2 f(\cdot)_d$  et par  $v(\cdot)$  la fonction  $v_{d,\varepsilon}(\cdot)$

A) Pour  $\theta \geq 0$ , désignons  $y + \theta d$  par  $y(\theta)$ .  
Choisissons arbitrairement  $x^*(\theta) \in M(y(\theta))$ ,  $x^* \in M(y)$   
et  $u(\theta) \in U(y(\theta))$ ,  $u \in U(y)$   
Puisque  $P_y$  et  $\varphi_y$  sont des problèmes duaux, il découle du lemme que  $(x^*, u)$  est un point selle de la fonction  $:(x^*, u) \rightarrow L(y, x^*, u)$ .  
Puisque  $x^*(\theta) \in M(y(\theta))$ ,  $u(\theta) \in U(y(\theta))$  et que

$$v(y(\theta)) = L(y(\theta), x^*(\theta), u(\theta)), \text{ on a } \forall x^* \forall u \geq 0 \\ L(y(\theta), x^*, u(\theta)) \leq L(y(\theta), x^*(\theta), u(\theta)) \leq L(y(\theta), x^*(\theta), u) \quad (1)$$

Si nous posons  $x^* = x^*(\theta)$  ds (1) et si nous nous rappelons que  $v(y) = L(y, x^*, u)$ , nous pouvons calculer  $\forall u \in U(y)$  que :

$$\frac{v(y(\theta)) - v(y)}{\theta} \leq \frac{1}{\theta} [L(y(\theta), x^*(\theta), u) - L(y, x^*, u)] \quad (2)$$

et le second membre de l'inégalité vaut

$$u [\langle x^*(\theta), \delta \rangle - \frac{f(y + \theta \delta) - f(y)}{\theta}] \quad (3)$$

B) Proverons maintenant la relation fondamentale (3.1.1)

a) Soit  $(\theta_m^1)_m$  une suite qui tend vers  $0^+$  telle que

$$\lim_{m \rightarrow \infty} \frac{v(y + \theta_m^1 \delta) - v(y)}{\theta_m^1} = \lim_{\theta \rightarrow 0^+} \sup_{\theta} \frac{v(y + \theta \delta) - v(y)}{\theta}$$

La limite supérieure existe car

$$\left| \frac{v(y + \theta \delta) - v(y)}{\theta} \right| \leq K \|\delta\| \text{ où } K \text{ est la constante de Lipschitz de } v(\cdot)$$

Puisque la multiapplication  $M(\cdot)$  est localement bornée, la suite  $\{x^*(\theta_m^1)\}_m$  a des points d'accumulation. Sans perdre de généralité, nous pouvons assurer que cette suite converge vers un point  $\bar{x}^*$ .

Puisque  $M(\cdot)$  est fermée,  $x^* \in M(y)$ .

Par (2) et (3), nous avons

$$\limsup_{\theta \rightarrow 0^+} \frac{v(y + \theta \delta) - v(y)}{\theta} \leq u[\langle x^*, \delta \rangle - f'(y; \delta)] \quad (4)$$

$\forall u \in U(y)$

b) Si nous posons  $u = u(\theta)$  dans (4), nous obtenons la minoration suivante:

$$\frac{v(y + \theta \delta) - v(y)}{\theta} \geq \frac{1}{\theta} [L(y(\theta), x^*, u(\theta)) - L(y, x^*, u(\theta))]$$

où le deuxième membre de l'inégalité est égal à

$$u(\theta) \left[ \langle x^*, \delta \rangle - \frac{f(y + \theta \delta) - f(y)}{\theta} \right]$$

Soit  $(\theta_m^2)_m$  une suite convergant vers  $0^+$  telle que

$$\lim_{m \rightarrow \infty} \frac{v(y + \theta_m^2 \delta) - v(y)}{\theta_m^2} = \liminf_{\theta \rightarrow 0^+} \frac{v(y + \theta \delta) - v(y)}{\theta}$$

Puisque  $U(\cdot)$  est localement bornée, la suite  $\{u(\theta_m^2)\}_m$  a des points d'accumulation. Sans perdre de généralité, nous pouvons assurer que  $\{u(\theta_m^2)\}_m$  converge vers  $\hat{u}$ . Puisque  $U(\cdot)$  est fermée,  $\hat{u} \in U(y)$  et nous obtenons :

$$\liminf_{\theta \rightarrow 0^+} \frac{v(y + \theta \delta) - v(y)}{\theta} \geq \hat{\mu} [\langle x^*, \delta \rangle - f'(y; \delta)] \quad \forall x^* \in M(y) \quad (5)$$

c) Si nous posons  $\mu = \hat{\mu}$  dans (4) et  $x^* = \hat{x}^*$  dans (5), il découle

$$v'(y; \delta) = \hat{\mu} [\langle \hat{x}^*, \delta \rangle - f'(y; \delta)]$$

De (4) et (5), on déduit alors

$$\max_{x^* \in M(y)} \hat{\mu} [\langle x^*, \delta \rangle - f'(y; \delta)] \leq v'(y; \delta) \leq \min_{\mu \in U(y)} \mu [\langle \hat{x}^*, \delta \rangle - f'(y; \delta)] \quad (6)$$

Utilisons maintenant le fait que  $\hat{x}^* \in M(y)$  et  $\hat{\mu} \in U(y)$  pour écrire :

$$\begin{aligned} \min_{\mu \in U(y)} \mu [\langle \hat{x}^*, \delta \rangle - f'(y; \delta)] &\leq \max_{x^* \in M(y)} \min_{\mu \in U(y)} \mu [\langle x^*, \delta \rangle - f'(y; \delta)] \\ &\leq \max_{x^* \in M(y)} \hat{\mu} [\langle x^*, \delta \rangle - f'(y; \delta)] \quad (7) \end{aligned}$$

Comme (7) est l'inégalité contraire de (6), il s'en suit :

$$v'(y; \delta) = \max_{x^* \in M(y)} \min_{\mu \in U(y)} \{ \mu [\langle x^*, \delta \rangle - f'(y; \delta)] \}$$

et les opérateurs max et min commutent.

□



Examinons quelques cas où  $v'_{d,\varepsilon}(x_0; \delta)$  a une expression simplifiée

① Si la forme linéaire  $x^* \mapsto \langle x^*, \delta \rangle$  est une constante  $\alpha$  sur  $\partial_\varepsilon f(x_0)_d$

c.à.d si  $\delta$  appartient au sous-espace orthogonal à l'enveloppe affine de  $\partial_\varepsilon f(x_0)_d$

Alors l'expression de  $v'_{\varepsilon,d}(x_0; \delta)$  devient :

$$v'_{\varepsilon,d}(x_0; \delta) = \min_{\mu \in M_{d,\varepsilon}(x_0)} \{ \mu [\alpha - f'(x_0; \delta)] \}$$

② Si  $\delta = p d$  où  $p$  est un réel, la forme linéaire

$x^* \mapsto \langle x^*, \delta \rangle$  est constante sur  $\partial_\varepsilon f(x_0)_d$ .

En effet,

$$\forall x^* \in \partial_\varepsilon f(x_0)_d \quad \langle x^*, d \rangle = f'_\varepsilon(x_0; d) = \beta.$$

$$\text{et donc } \langle x^*, \delta \rangle = \langle x^*, p d \rangle = p \beta.$$

Dès lors,

$$v'_{d,\varepsilon}(x_0; p d) = \min_{\mu \in M_{d,\varepsilon}(x_0)} \{ \mu p [\beta - f'(x_0; d)] \}$$



La situation particulière où  $\delta = d$  est appelée cas diagonal et nous avons :

$$v'_{d,\varepsilon}(x_0; d) = f''_{\varepsilon}(x_0; d, d) = \min_{\gamma \in M_{d,\varepsilon}(x_0)} \left\{ \gamma[f'_{\varepsilon}(x_0; d) - f'(x_0; d)] \right\}$$

③

Si  $\partial_{\varepsilon} f(x_0)_d$  est un singleton  $\{x_d^*\}$

Alors

$$v'_{d,\varepsilon}(x_0; \delta) = \min_{\gamma \in M_{d,\varepsilon}(x_0)} \left\{ \gamma[\langle x_d^*, \delta \rangle - f'(x_0; \delta)] \right\}$$

Dans ce dernier cas, il est possible de donner une expression du sous-différentiel de la fonction  $v_{d,\varepsilon}(\cdot)$ . Ceci sera développé au paragraphe suivant.

### III.2 Sous-différentiabilité de $v_{d,\varepsilon}(\cdot)$

Le caractère localement lipschitz de la fonction  $v_{d,\varepsilon}(\cdot)$  (voir II.4) nous permet de définir son gradient généralisé au sens de Clarke.

Rappelons que celui-ci, noté  $\partial v_{d,\varepsilon}(x_0)$ , est l'enveloppe convexe de l'ensemble

$$\left\{ x^* \text{ t.q. } \exists (x_k)_k \longrightarrow x_0 \text{ dans } E_1 \text{ avec } x_k \in \mathbb{R}^n \text{ et } \nabla v_{d,\varepsilon}(x_k) \rightarrow x^* \right\}$$

où  $E_1$  est l'ensemble des points  $x$  où  $v_{d,\varepsilon}$  est différentiable

.  $\mathbb{R}^n \setminus E_1$  est un ensemble de mesure nulle.

La fonction d'appui de  $\partial v_{d,\varepsilon}(x_0)$  est la dérivée directionnelle généralisée, notée  $v_{d,\varepsilon}^0(x_0; \delta)$ .

$$v_{d,\varepsilon}^0(x_0; \delta) = \lim_{\substack{x \rightarrow x_0 \\ \lambda \rightarrow 0}} \sup \frac{v_{d,\varepsilon}(x + \lambda \delta) - v_{d,\varepsilon}(x)}{\lambda}$$

$$= \lim_{\substack{x \rightarrow x_0 \\ x \in E_1}} \sup \langle \nabla v_{d,\varepsilon}(x), \delta \rangle$$

Par [Chap I (5.4)], nous savons que, si la dérivée directionnelle au sens habituel dans la direction  $\delta$

existe, alors nous avons

$$v'_{d,\varepsilon}(x_0; \delta) \leq v^0_{d,\varepsilon}(x_0; \delta).$$

Une première expression de  $\partial v_{d,\varepsilon}(x_0)$  est donnée par la proposition suivante, dont la démonstration se trouve en ([1], p.37)

Proposition 3.2.1.

Supposons que  $\partial_\varepsilon f(x)_d$  est un singleton  $\{x_d^*\}$  au voisinage de  $x_0$

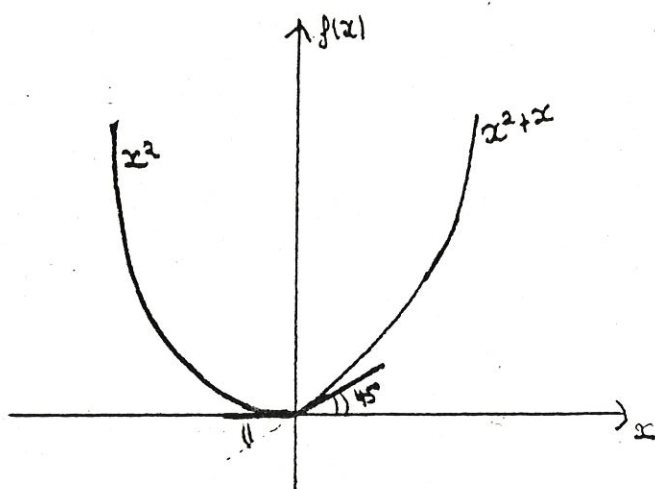
$$\text{Alors } \partial v_{d,\varepsilon}(x_0) = M_{d,\varepsilon}(x_0) [x_d^* - \partial f(x_0)]$$

Illustrons cette proposition par un exemple où la fonction  $f$  est quadratique par morceaux (exemple 5).

Soit  $f : \mathbb{R} \longrightarrow \mathbb{R}$

$$x \longmapsto \begin{cases} x^2 & \text{si } x \leq 0 \\ x^2 + x & \text{si } x \geq 0 \end{cases}$$

Prends  $x_0 = 0$  et  $d = 1$



Le calcul du sous-différentiel de  $V_{d,\varepsilon}(\cdot)$  au point  $x_0$  nécessite la connaissance de  $\partial f(x_0)$ , de  $x_d^*$  et de  $M_{d,\varepsilon}(x_0)$ .

a) calcul de  $\partial f(x_0)$

Comme  $\partial f(x_0) = \{x^* \in \mathbb{R} \mid f(x) \geq f(x_0) + (x - x_0)x^* \forall x \in \mathbb{R}\}$

on a dans cet exemple.

$$f(x) \geq x x^* \quad \forall x \in \mathbb{R}.$$

$$\cdot \text{ si } x > 0,$$

$$x^2 + x \geq x x^*$$

$$\text{d'où } x^* \leq 1.$$

$$\cdot \text{ si } x < 0$$

$$x^2 \geq x x^*.$$

$$\text{et donc } x^* \geq 0.$$

On tire alors  $\partial f(0) = [0, 1]$ , ce qui peut se vérifier graphiquement : les droites passant par 0 ayant un coefficient angulaire compris entre 0 et 1

ont un épigraphe qui contient l'épigraphe de  $f$ .  
 la fonction d'appui de  $\partial f(0) = f'(0; 1) = 1$

b) Calcul de  $M_{d,\varepsilon}(x_0)$  et de  $f'_\varepsilon(x_0; d)$ .

Recherchons les points qui minimisent  $q_{d,\varepsilon}(\cdot)$ .

$$q_{d,\varepsilon}(\lambda) = \frac{f(\lambda) + \varepsilon}{\lambda} = \lambda + 1 + \frac{\varepsilon}{\lambda}$$

Cette fonction atteint son minimum  $2\sqrt{\varepsilon} + 1$  en  
 $\lambda = \sqrt{\varepsilon}$

$$\text{Dès lors } \Lambda_{d,\varepsilon}(x_0) = \{\sqrt{\varepsilon}\}$$

$$M_{d,\varepsilon}(x_0) = \left\{ \frac{1}{\sqrt{\varepsilon}} \right\}$$

$$f'_\varepsilon(x_0; d) = 2\sqrt{\varepsilon} + 1$$

c) Calcul de  $\{x_d^*\} = \partial_\varepsilon f(x_0)_d$

$$\partial_\varepsilon f(x_0)_d = \left\{ x^* \in \partial_\varepsilon f(x_0) \text{ t.q. } \langle x^*, d \rangle = f'_\varepsilon(x_0; d) \right\}$$

Calculons donc d'abord  $\partial_\varepsilon f(x_0)$

$$\text{Comme } \partial_\varepsilon f(x_0) = \left\{ x^* \in \mathbb{R} \text{ t.q. } f(x) \geq f(x_0) + (x - x_0, x^*) - \varepsilon \quad \forall x \in \mathbb{R} \right\}$$

$$\text{on a } f(x) \geq xx^* - \varepsilon \quad \forall x \in \mathbb{R}.$$

$$\bullet \text{ Si } x > 0 \text{ alors } x^* \leq x + 1 + \frac{\varepsilon}{x}$$

$$\text{Donc } x^* \leq 2\sqrt{\varepsilon} + 1$$

$$\bullet \text{ Si } x < 0 \text{ alors } x^* \geq x + \frac{\varepsilon}{x}.$$

La valeur maximale de la fonction  $x + \frac{\varepsilon}{x}$  sur  $\mathbb{R}^-$



est  $-2\sqrt{\varepsilon}$ .

Donc  $x^* \geq -2\sqrt{\varepsilon}$

Des lors  $\partial_{\varepsilon} f(x_0) = [-2\sqrt{\varepsilon}, 2\sqrt{\varepsilon}+1]$  et  $\partial_{\varepsilon} f(x_0)_d = \{2\sqrt{\varepsilon}+1\}$

d) Nous sommes maintenant en mesure de trouver l'expression du sous-différentiel de  $v_{d,\varepsilon}(\cdot)$  en 0

$$\partial v_{d,\varepsilon}(0) = M_{d,\varepsilon}(0) [x_d^* - \partial f(0)]$$

$$= \left[ 2, 2 + \frac{1}{\sqrt{\varepsilon}} \right]$$

Remarquons que, si l'on calcule la fonction d'appui du sous-différentiel de Clarke pour la direction  $\delta = d=1$ , on obtient

$$v_{d,\varepsilon}^0(x_0; d) = 2 + \frac{1}{\sqrt{\varepsilon}}$$

tandis que la dérivée directionnelle au sens habituel vaut

$$v'_{d,\varepsilon}(x_0; d) = \min_{\eta \in M_{d,\varepsilon}(x_0)} [\eta(f'_{\varepsilon}(x_0; d) - f'(x_0; d))] = 2$$

et sa valeur est inférieure à  $v_{d,\varepsilon}^0(x_0; d)$

On peut particulariser la thèse de la proposition précédente en émettant une hypothèse plus restrictive sur la fonction  $f$ .

Proposition 3.2.2.

Soit  $f$  différentiable sur  $\mathbb{R}^n$

Alors  $\forall x \in \mathbb{R}^n, \partial V_{d,\varepsilon}(x) = M_{d,\varepsilon}(x) [x_d^*(x) - \nabla f(x)]$

où  $\partial_\varepsilon f(x)_d = \{x_d^*(x)\}$

La proposition 3.2.2 nous suggère que, sous cette même hypothèse sur  $f$ , si  $M_{d,\varepsilon}(x)$  est un singleton,  $\partial V_{d,\varepsilon}(x)$  sera lui aussi réduit à un singleton. Cette condition est satisfaite, par exemple, si  $f$  est strictement convexe.

Proposition 3.2.3

Soit  $f$  différentiable et strictement convexe sur  $\mathbb{R}^n$

Alors  $V_{d,\varepsilon}$  est continûment différentiable et

$$\nabla V_{d,\varepsilon}(x) = \mu_d(x) [x_d^*(x) - \nabla f(x)]$$

où  $\{\mu_d(x)\} = M_{d,\varepsilon}(x)$

Dans les trois propositions dont les démonstrations se trouvent en ([1], p.38), une des difficultés pour calculer  $\partial V_{d,\varepsilon}(x)$  est de trouver  $\partial_\varepsilon f(x)_d$ .

La proposition suivante donne une simplification de son expression.

Proposition 3.2.4

Si  $\Lambda_{d,\varepsilon}(x_0)$  est non vide

$$\text{Alors } \partial_{\varepsilon} f(x_0)_d = \{x^* \in \partial f(x_0 + \lambda d) \text{ t.q. } \langle x^*, d \rangle = f'_{\varepsilon}(x_0; d)\} \\ \forall \lambda \in \Lambda_{d,\varepsilon}(x_0)$$

Énonçons quelques remarques concernant  $\partial_{\varepsilon} f(x)_d$

Remarque 1

Nous prouvons en annexe 4 que  $\partial_{\varepsilon} f(x)_d$  est le sous-différentiel de la fonction  $f'_{\varepsilon}(x_0; d)$  en la variable  $d$ . (\*)  
Or celle-ci est convexe et donc différentiable presque partout [chap 1 (4.7)]

Dès lors  $\partial_{\varepsilon} f(x)_d$  est un singleton pour presque toute direction  $d$

Cependant, il y a des situations où  $\partial_{\varepsilon} f(x)_d$  est un singleton pour tout  $d$ . Il en est ainsi pour une fonction définie sur la droite réelle. En effet, dans ce cas, il n'y a que deux directions  $-1$  et  $+1$ . L'ensemble  $\partial_{\varepsilon} f(x)_d$  est alors réduit à  $\{x_d^*\}$  où  $x_d^*$  est la borne supérieure

ou inférieure de l'intervalle fermé  $\partial_\varepsilon f(x)$ .

Dans la suite, nous noterons parfois l'ensemble  $\partial_\varepsilon f(x_0)_d$  par  $\partial \psi_{\partial_\varepsilon f(x_0)}^*(d)$  où  $\psi_A^*$  représente la fonction d'appui de  $A$ . Ceci est légitimé par (\*).

### Remarque 2

Si  $f$  est différentiable et si une direction telle que  $\Lambda_{d,\varepsilon}(x) \neq \emptyset$ , il découle de la proposition 3.2.4 que  $\partial_\varepsilon f(x)_d = \{ \nabla f(x + \lambda d) \} \quad \forall \lambda \in \Lambda_{d,\varepsilon}(x)$ .

### Remarque 3

Géométriquement,  $\partial_\varepsilon f(x_0)_d$  peut s'interpréter comme l'ensemble de tous les points  $x^* \in \partial_\varepsilon f(x_0)$  pour lesquels  $d$  est un vecteur normal à l'hyperplan engendré par  $\partial_\varepsilon f(x_0)$ .

Le calcul de  $\partial_\varepsilon f(x)_d$ , plus aisé grâce à la proposition 3.2.4 nous permet d'illustrer maintenant la proposition 3.2.3.

### Exemple 6

Soit  $f : \mathbb{R}^n \longrightarrow \mathbb{R}$

$$x \longmapsto \frac{1}{2} \langle Ax, x \rangle + \langle b, x \rangle + c$$

où  $A$  est une matrice  $(n \times n)$  symétrique  
définie positive



$b$  un vecteur de  $\mathbb{R}^n$

$\varepsilon \in \mathbb{R}$

Cette fonction quadratique satisfait les hypothèses de la proposition 3.2.3, à savoir, la différentiabilité et la stricte convexité sur  $\mathbb{R}^n$ .

Nous avons déjà calculé dans le chap II (ex 3) que

$$\lambda_{d,\varepsilon}(x_0) = \left\{ \left( \frac{2\varepsilon}{\langle Ad, d \rangle} \right)^{1/2} \right\}, \text{ et donc } \mu_{d,\varepsilon}(x_0) = \left\{ \left( \frac{\langle Ad, d \rangle}{2\varepsilon} \right)^{1/2} \right\}$$

Par la remarque 2, nous avons

$$\partial_\varepsilon f(x)_d = \left\{ \nabla f(x_0 + \lambda d) \right\} \text{ où } \lambda \in \lambda_{d,\varepsilon}(x_0).$$

$$\text{Dès lors } \partial_\varepsilon f(x)_d = \left\{ x_d^*(x_0) \right\}$$

$$\text{où } x_d^*(x_0) = Ax_0 + b + \left( \frac{2\varepsilon}{\langle Ad, d \rangle} \right)^{1/2} Ad$$

La proposition 3.2.3 nous renseigne que

$$\nabla v_{d,\varepsilon}(x_0) = \mu_{d,\varepsilon}(x_0) [x_d^*(x_0) - \nabla f(x_0)]$$

$$\text{Dès lors } \nabla v_{d,\varepsilon}(x_0) = Ad \text{ et}$$

$$v'_{\varepsilon,d}(x_0; \delta) = j'_\varepsilon(x_0; d, \delta) = \langle \nabla v_{d,\varepsilon}(x_0), \delta \rangle = \langle Ad, \delta \rangle.$$

Remarquons que, dans cet exemple,  $\mu_{d,\varepsilon}(x_0) \rightarrow +\infty$  et  $[x_d^*(x_0) - \nabla f(x_0)] \rightarrow 0$  quand  $\varepsilon \rightarrow 0$ . Donc, même sous des hypothèses assez fortes sur la fonction  $f$ , le comportement du gradient de  $v_{\varepsilon,d}(\cdot)$  peut être inattendu quand  $\varepsilon \rightarrow 0$ . Ceci fera l'objet du chapitre suivant.



Auparavant, nous poursuivons l'étude de la différentiabilité de  $v_{d,\varepsilon}(\cdot)$  en élargissant les hypothèses des propositions précédentes, à savoir, la non-réduction de  $\partial_\varepsilon f(x)_d$  à un singleton.

### Théorème 3.2

Soit  $x$  un point où  $v_{d,\varepsilon}$  est différentiable

Alors i)  $\eta_{d,\varepsilon}(x)$  est un singleton  $\{\eta_{d,\varepsilon}(x)\}$  et

$$\eta_{d,\varepsilon}(x) \partial_\varepsilon f(x)_d = \nabla v_d(x) + \eta_{d,\varepsilon}(x) \partial f(x)$$

ii)  $\nabla v_{d,\varepsilon}(x) \neq 0$ ssi  $\eta_{d,\varepsilon}(x) > 0$ .

La démonstration de ce théorème nécessite le lemme suivant.

### Lemme 3.2

Supposons que  $\eta_{d,\varepsilon}(x)$  est non vide

Alors  $f'_\varepsilon(x;d) > f'(x;d)$ .

### Preuve du lemme

D'après la définition de  $\eta_{d,\varepsilon}(x)$ , on a pour tout  $\lambda_0 \in \eta_{d,\varepsilon}(x)$

$$\frac{f(x + \lambda_0 d) - f(x) + \varepsilon}{\lambda_0} = f'_\varepsilon(x;d)$$

tandis que  $\inf_{\lambda > 0} \frac{f(x + \lambda d) - f(x)}{\lambda} = f'(x;d)$

et donc nous avons l'inégalité stricte annoncée  
puisque  $\varepsilon > 0$ .

### Preuve du théorème

Soit  $x$  un point où  $v_{d,\varepsilon}(\cdot)$  est différentiable

i) Pour toute direction  $\delta \in \mathbb{R}^n$ , on a d'après le théorème 3.1

$$\langle \nabla v_{d,\varepsilon}(x), \delta \rangle = \min_{\mu \in M_{d,\varepsilon}(x_0)} \max_{x^* \in \partial_\varepsilon f(x_0)_d} \{ \mu [ \langle x^*, \delta \rangle - f'(x; \delta) ] \} \quad (3.2.1)$$

Si on réécrit  $\langle \nabla v_{d,\varepsilon}(x), \delta \rangle = -\langle \nabla v_{d,\varepsilon}(x), -\delta \rangle$ , la relation (3.2.1) donne

$$\begin{aligned} \min_{\mu \in M_{d,\varepsilon}(x_0)} \max_{x^* \in \partial_\varepsilon f(x_0)_d} \{ \mu [ \langle x^*, \delta \rangle - f'(x; \delta) ] \} &= \langle \nabla v_{d,\varepsilon}(x), \delta \rangle \\ &= \max_{\mu \in M_{d,\varepsilon}(x_0)} \min_{x^* \in \partial_\varepsilon f(x_0)_d} \{ \mu [ \langle x^*, \delta \rangle + f'(x; -\delta) ] \} \quad (3.2.2) \end{aligned}$$

Faisons  $\delta = d$  dans (3.2.2) pour obtenir

$$\min_{\mu \in M_{d,\varepsilon}(x_0)} \{ \mu [ f'_\varepsilon(x; d) - f'(x; d) ] \} = \max_{\mu \in M_{d,\varepsilon}(x_0)} \{ \mu [ f'_\varepsilon(x; d) + f'(x; -d) ] \} \quad (3.2.3)$$

$$\text{OR } -f'(x; d) \leq f'(x; -d)$$

$$\text{car } f'(x; -d) = \max_{x^* \in \partial f(x)} \langle x^*, -d \rangle = -\min_{x^* \in \partial f(x)} \langle x^*, d \rangle$$

$$\geq -\max_{x^* \in \partial f(x)} \langle x^*, d \rangle = -f'(x; d)$$

Ceci nous permet de tirer de (3.2.3) que

$$\begin{aligned} \max_{\gamma \in M_{d,\varepsilon}(x_0)} \left\{ \gamma [f'_\varepsilon(x;d) - f'(x;d)] \right\} &\leq \max_{\gamma \in M_{d,\varepsilon}(x_0)} \left\{ \gamma [f'_\varepsilon(x;d) + f'(x;-d)] \right\} \\ &\leq \min_{\gamma \in M_{d,\varepsilon}(x_0)} \left\{ \gamma [f'_\varepsilon(x;d) - f'(x;d)] \right\} \end{aligned}$$

Comme l'inégalité contraire est toujours vraie, on obtient

$$\max_{\gamma \in M_{d,\varepsilon}(x_0)} \left\{ \gamma [f'_\varepsilon(x;d) - f'(x;d)] \right\} = \min_{\gamma \in M_{d,\varepsilon}(x_0)} \left\{ \gamma [f'_\varepsilon(x;d) - f'(x;d)] \right\} \quad (3.2.4)$$

On sait que  $f'_\varepsilon(x;d) \geq f'(x;d) \quad \forall \varepsilon \geq 0$

Examinons dès lors deux cas, suivant que l'inégalité est stricte ou non

- si  $f'_\varepsilon(x;d) > f'(x;d)$ .

alors il découle de (3.2.4) que  $M_{d,\varepsilon}(x) = \{\gamma_{d,\varepsilon}(x)\}$ .

- si  $f'_\varepsilon(x;d) = f'(x;d)$

alors le lemme (3.2) nous apprend que  $M_{d,\varepsilon}(x)$  est vide et donc  $M_{d,\varepsilon}(x)$  est réduit au singleton  $\{0\}$

Dans les deux cas, l'ensemble  $M_{d,\varepsilon}(x)$  est un singleton.

Pourrons maintenant la seconde assertion de i), c.à.d

$$\gamma_{d,\varepsilon}(x) \partial_\varepsilon f(x)_d = \nabla v_{d,\varepsilon}(x) + \gamma_{d,\varepsilon}(x) \partial f(x).$$

Puisque  $M_{d,\varepsilon}(x)$  est un singleton, (3.2.1) se réécrit comme suit, pour toute direction  $d$

$$\max_{x^* \in \partial_\varepsilon f(x)_d} [\mu_d(x) \langle x^*, \delta \rangle] = \langle \nabla v_{d,\varepsilon}(x), \delta \rangle + \mu_{d,\varepsilon}(x) f'(x; d)$$

Comme les deux membres de cette égalité sont respectivement des fonctions d'appui des ensembles  $\{\mu_{d,\varepsilon}(x) \partial_\varepsilon f(x)_d\}$  et  $\{\nabla v_{d,\varepsilon}(x) + \mu_{d,\varepsilon}(x) \partial f(x)\}$ , ceux-ci sont égaux.

ii) Montrons que  $\nabla v_{d,\varepsilon}(x) = 0$  si  $\mu_{d,\varepsilon}(x) = 0$ .

. la condition est suffisante

si  $\mu_{d,\varepsilon}(x) = 0$  alors  $\nabla v_{d,\varepsilon}(x) = 0$  par i)

. la condition est nécessaire

si  $\nabla v_{d,\varepsilon}(x) = 0$ , on a par i).

$$\mu_{d,\varepsilon}(x) \partial_\varepsilon f(x)_d = \mu_{d,\varepsilon}(x) \partial f(x).$$

Si  $\mu_{d,\varepsilon}(x)$  n'était pas nul, cette égalité se réduirait à  $\partial_\varepsilon f(x)_d = \partial f(x)$ .

Ceci conduirait à l'égalité des fonctions d'appui  $f'_\varepsilon(x; d)$  et  $f'(x; d)$ . Donc par le lemme, nous pourrions affirmer que  $M_{d,\varepsilon}(x_0) = \emptyset$ , ou encore que  $M_{d,\varepsilon}(x_0) = \{0\}$ . Ceci est en contradiction avec la non-nullité de  $\mu_{d,\varepsilon}(x)$ .

□



### Conséquence du théorème 2.2

Sous l'hypothèse de la différentiabilité de  $v_{d,\varepsilon}(\cdot)$  au point  $x$ , assurer que  $\partial_\varepsilon f(x)_d$  est un singleton  $x_d^*(x)$  est équivalent à assurer la différentiabilité de  $f$ .

Dans un tel cas, on a la relation

$$x_d^*(x) = \eta_{d,\varepsilon}^{-1} \nabla v_{d,\varepsilon}(x) + \nabla f(x).$$

Les résultats de ce paragraphe peuvent être précisés si nous recueillons plus de renseignements sur l'ensemble  $\Lambda_{d,\varepsilon}(x)$ . Par exemple, si l'on est assuré que  $\Lambda_{d,\varepsilon}(x)$  est non vide, on pourra prouver que le gradient de  $v_{d,\varepsilon}(x)$  n'est jamais nul. C'est ce genre de situations qui va maintenant capter notre attention.



### III.3 Caractérisations de l'ensemble $\Lambda_{d,\varepsilon}(x_0)$ et conséquences sur l'expression du gradient de $v_{d,\varepsilon}$ .

---

Comme suggéré dans le second chapitre, lors de l'illustration des propriétés de la fonction  $q_{d,\varepsilon}(\lambda)$  (exemple 4), l'ensemble des minimas de cette fonction peut être non vide et borné, non borné ou vide.

Nous allons approfondir l'étude de ces situations.

Auparavant, remarquons que cet ensemble  $\Lambda_{d,\varepsilon}(x_0)$  ne peut être la réunion de plusieurs intervalles disjoints vu la pseudo-convexité de  $q_{d,\varepsilon}(\lambda)$ .

Premier cas :  $\Lambda_{d,\varepsilon}(x_0)$  est non vide et borné.

---

Le cas où  $\Lambda_{d,\varepsilon}(x_0)$  est non vide et borné correspond à la situation où il existe  $\lambda^*$  tel que

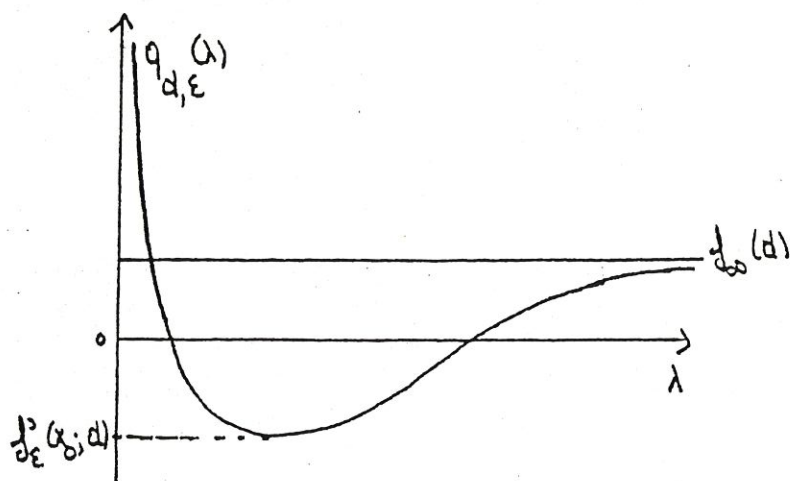
$$\boxed{S_1}$$

$$q_{d,\varepsilon}(\lambda^*) = \frac{f(x_0 + \lambda^* d) - f(x_0) + \varepsilon}{\lambda^*} < f'_\infty(d)$$

$$\text{où } f'_\infty(d) = \lim_{\lambda \rightarrow +\infty} q_{d,\varepsilon}(\lambda) = \sup_{\lambda > 0} \frac{f(x_0 + \lambda d) - f(x_0)}{\lambda}$$

Cette situation, notée  $S_1$ , se réalise quand  $q_{d,\varepsilon}$  atteint son minimum en dehors de  $f'_\infty(d)$ .

Elle se représente comme suit :



Rappelons que  $M_{d,\epsilon}(x_0) = \left\{ \frac{1}{\lambda} \mid \lambda \in \Lambda_{d,\epsilon}(x_0) \right\}$  et donc, dans ce cas,  $M_{d,\epsilon}(x_0)$  est un intervalle compact ne contenant pas 0. Ceci nous permet de préciser le résultat du théorème 3.2 sur la différentiabilité de  $v_{d,\epsilon}(z)$ .

### Théorème 3.3.

Soit  $d$  une direction satisfaisant  $(S_1)$  en  $x_0$ .

Alors il existe un voisinage  $V$  de  $x_0$  tel que

pour chaque  $x \in V$ ,

si  $v_{d,\epsilon}$  est différentiable en  $x$

alors  $\eta_{d,\epsilon}(x) > 0$  et

$$\partial_\epsilon f(x)_d = \eta_{d,\epsilon}(x)^{-1} \nabla v_{d,\epsilon}(x) + \partial f(x) \quad (3.3.1)$$

### Démonstration

Par la semi-continuité supérieure de la multiapplication  $x \rightrightarrows M_{d,\varepsilon}(x)$ , l'ensemble  $M_{d,\varepsilon}(x)$  est inclus dans  $R_0^+$  pour tout  $x$  du voisinage  $V$  de  $x_0$ .

Le théorème 3.2 nous assure que  $M_{d,\varepsilon}(x)$  est réduit à un singleton  $\mu_{d,\varepsilon}(x)$ . Il sera donc strictement positif chaque fois que  $v_{d,\varepsilon}$  est différentiable en  $x \in V$ .

L'inégalité (3.31) résulte du théorème 3.2.

□

Comme l'ensemble  $\partial_\varepsilon f(x)_d$  est presque partout un singleton  $\{x_d^*(x)\}$ , il est intéressant d'énoncer le corollaire suivant :

### Corollaire 1

Soit  $d$  satisfaisant  $(S_1)$  en  $x_0$ .

Alors dans un voisinage  $V$  de  $x_0$ , les assertions suivantes sont équivalentes :

- a)  $v_{d,\varepsilon}$  et  $f$  sont différentiables en  $x$
- b)  $M_{d,\varepsilon}(x) = \{\mu_{d,\varepsilon}(x)\}$  et  $\partial_\varepsilon f(x)_d = \{x_d^*(x)\}$

En de tels points  $x$  on a :

$$\nabla v_{d,\varepsilon}(x) = \mu_{d,\varepsilon}(x) [x_d^*(x) - \nabla f(x)]$$

Avant d'illustrer cette situation, énonçons trois remarques que nous pourrions également valider par un exemple.

Remarque 1

$$\forall x \in V \quad v'_{d,\varepsilon}(x;d) > 0$$

. Notons que  $\nabla v'_{d,\varepsilon}(x) \neq 0$  chaque fois que  $v'_{d,\varepsilon}(x)$  est différentiable en  $x$

. Comme  $\forall x \in V \quad M_{d,\varepsilon}(x) \subset \mathbb{R}_0^+$ ,  $v'_{d,\varepsilon}(x)$  est certainement mon croissant. Par le lemme 3.2, nous avons donc

$$f'_\varepsilon(x;d) > f'(x;d) \quad (3.3.2)$$

De plus  $v''_{d,\varepsilon}(x;d)$  est une matrice  $f''_\varepsilon(x;d;d)$  vaut

$$\min \left\{ \eta \left[ f'_\varepsilon(x;d) - f'(x;d) \right] \right\}$$

$$\eta \in M_{d,\varepsilon}(x)$$

Par l'inégalité (3.3.2) et le fait que  $M_{d,\varepsilon}(x) \subset \mathbb{R}_0^+$ , nous obtenons que

$$f''_\varepsilon(x;d;d) > 0 \quad \forall x \in V.$$

□

Remarque 2 : La différentiabilité de  $v'_{d,\varepsilon}$  n'implique pas

nécessairement celle de  $f$

Mais néanmoins nous avons la relation suivante :

$$\forall \delta \text{ t.q. } (x^\delta, \delta) \text{ est constante sur } \partial_\varepsilon f(x)_d$$

$$f'(x;d) = -f'(x;-\delta)$$



En effet

$$f'(x; \delta) = \max \langle x^*, \delta \rangle$$

$$x^* \in \partial f(z)$$

$$= \max \langle x^*, \delta \rangle$$

$$x^* \in \left[ \partial_\varepsilon f(z)_D - (\eta_{d,\varepsilon}(z))^{-1} \nabla v_{d,\varepsilon}(z) \right]$$

par le théorème 3.2

$$= \max \langle x^*, \delta \rangle + \langle -\eta_{d,\varepsilon}(z)^{-1} \nabla v_{d,\varepsilon}(z), \delta \rangle$$

$$x^* \in \partial_\varepsilon f(z)_D$$

$$= \alpha + \langle -\eta_{d,\varepsilon}(z)^{-1} \nabla v_{d,\varepsilon}(z), \delta \rangle$$

De la même manière,  $f'(z; -\delta) = -\alpha + \langle -\eta_{d,\varepsilon}(z)^{-1} \nabla v_{d,\varepsilon}(z), -\delta \rangle$

$$\text{Donc } f'(x; \delta) = -f'(x; -\delta)$$

□

### Remarque 3

La situation  $(S_1)$  est automatiquement réalisée lorsque  $f_0(d) = +\infty$  puisque  $q_{d,\varepsilon}(x)$  prend au moins une valeur finie.

Illustrons maintenant le théorème 3.3 par un exemple

### Exemple 7

$$\text{Soit } f: \mathbb{R}^2 \longrightarrow \mathbb{R}$$

$$(x, y) \longmapsto \begin{cases} 0 & \text{si } y < 0 \\ x^2 + y & \text{si } y \geq 0 \end{cases}$$

On considère le point  $x_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ , où  $f$  n'est pas différentiable.

Choisissons une direction  $d$  quelconque dans le premier quadrant c.-à-d.  $d = \begin{pmatrix} \alpha \\ 1-\alpha \end{pmatrix}$  avec  $0 < \alpha \leq 1$ .

Cet exemple concrétise la situation (S<sub>1</sub>) car  $f_\infty(d) = +\infty$

En effet

$$q_d(\lambda) = \frac{f(x_0 + \lambda d) - f(x_0)}{\lambda} = \lambda \alpha^2 + (1-\alpha).$$

$$\text{et } \lim_{\lambda \rightarrow +\infty} q_d(\lambda) = +\infty$$

Pour vérifier que  $f''_\varepsilon(x_0; d, d)$  est  $> 0$ , calculons successivement  $f'(x_0; d)$ ,  $M_{d,\varepsilon}(x_0)$ ,  $f'_\varepsilon(x_0; d)$

$$\bullet f'(x_0; d) = \inf_{\lambda > 0} q_d(\lambda) = 1 - \alpha$$

$$\bullet q_{d,\varepsilon}(\lambda) = \lambda \alpha^2 + (1-\alpha) + \frac{\varepsilon}{\lambda}.$$

Cette fonction atteint son minimum  $2\sqrt{\varepsilon}\alpha + (1-\alpha)$  en  $\lambda = \frac{\sqrt{\varepsilon}}{\alpha}$ .

$$\text{Des lors } \Lambda_{d,\varepsilon}(x_0) = \left\{ \frac{\sqrt{\varepsilon}}{\alpha} \right\}, \quad M_{d,\varepsilon}(x_0) = \left\{ \frac{\alpha}{\sqrt{\varepsilon}} \right\} \text{ et}$$

$$f'_\varepsilon(x_0; d) = 2\sqrt{\varepsilon}\alpha + (1-\alpha)$$

$$\bullet f''_\varepsilon(x_0; d, d) = \frac{\alpha}{\sqrt{\varepsilon}} [2\sqrt{\varepsilon}\alpha + (1-\alpha) - (1-\alpha)] = 2\alpha^2$$

on vérifie bien que

$$f''_\varepsilon(x_0; d, d) > 0 \quad \forall \alpha > 0$$

### Remarque concernant l'exemple

Nous avons écarté la possibilité  $\alpha = 0$  (c.-à-d.  $d = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ ) car ce cas dépasse le cadre de la situation ( $S_1$ ).

En effet  $q_{d,\varepsilon}(\lambda) = 1 + \frac{\varepsilon}{\lambda} > f_\infty(d) = 1 \quad \forall \lambda$

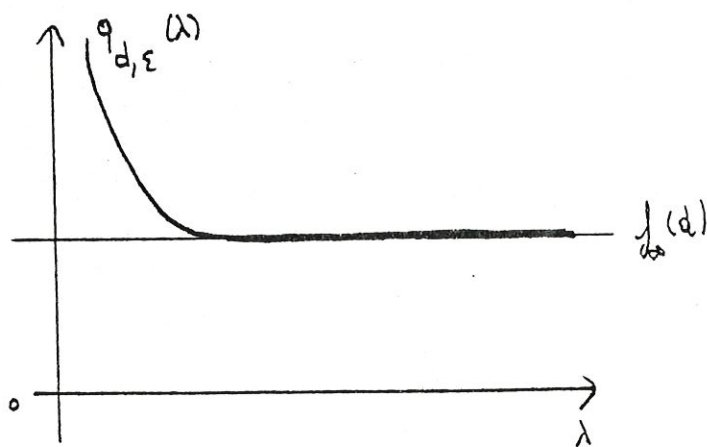
Ceci anticipe sur le troisième cas. Auparavant, considérons  $\lambda_{d,\varepsilon}(x_0)$  quand il est non vide et non borné.

Deuxième cas :  $\lambda_{d,\varepsilon}(x_0)$  est non vide et non borné

Ceci signifie que la fonction  $q_{d,\varepsilon}(\lambda)$  atteint sa valeur minimale sur la droite  $y = f_\infty(d)$ , c.-à-d.

$$\boxed{S_2} \quad \min_{\lambda > 0} \frac{f(x_0 + \lambda d) - f(x_0) + \varepsilon}{\lambda} = f_\infty(d)$$

Schématisons cette situation, notée ( $S_2$ ).



Signalons d'une part que, si la fonction  $q_{d,\varepsilon}(\lambda)$  atteint

son minimum sur  $f_0(d)$ , alors l'ensemble  $\Lambda_{d,\varepsilon}(x_0)$  est nécessairement non borné puisque cette fonction est quasi-convexe et que sa limite à l'infini vaut  $f_0(d)$ . D'autre part, puisque  $\Lambda_{d,\varepsilon}(x_0)$  est non vide et non borné,  $M_{d,\varepsilon}(x_0)$  contient 0 mais n'est pas réduit à cet élément. Introduisons à présent deux nouvelles notations, à savoir  $\underline{\lambda}_\varepsilon$  et  $\overline{\lambda}_\varepsilon$ . Grâce une expression simplifiée de  $M_{d,\varepsilon}(x_0)$ , elle permettant une écriture plus aisée de  $f'_\varepsilon(x_0; d)$  et  $f''_\varepsilon(x_0; d)$ . Quand  $\Lambda_{d,\varepsilon}(x_0)$  est non vide, nous notons :

$$\underline{\lambda}_\varepsilon = \min \Lambda_{d,\varepsilon}(x_0)$$

$$\overline{\lambda}_\varepsilon = \max \Lambda_{d,\varepsilon}(x_0)$$

Dans  $(S_2)$ ,  $\overline{\lambda}_\varepsilon = +\infty$ . On peut dès lors écrire

$$M_{d,\varepsilon}(x_0) = \left[ 0, \frac{1}{\underline{\lambda}_\varepsilon} \right].$$

Comme dans la situation  $(S_1)$ , particularisons le théorème 3.2 par un énoncé. (dont la démonstration se trouve en ([2], p.14))

#### Théorème 3.4.

Soit  $d$  satisfaisant  $(S_2)$  en  $x_0$

Soit  $V$  voisinage ouvert de  $x_0$

$$V_0 = \{x \in V \text{ où } V_{0,\varepsilon} \text{ est différentiable}\}$$

Alors  $\nabla V_{d,\varepsilon}(x)$  est non nul sur un ensemble  $W$  de mesure positive de  $V_0$ .



Précédemment, nous avons mis en évidence que  $x_0$  satisfait  $(S_1)$  en  $x_0$ , alors il existe un voisinage  $V$  de  $x_0$  où  $f''_{\varepsilon}(x; d, d) > 0 \forall x \in V$ . Par contre,  $x_0$  satisfait  $(S_2)$  en  $x_0$ , alors on peut prouver que  $f''_{\varepsilon}(x; d, d)$  est strictement positif presque partout sur  $W$  ([2], p.14) mais certainement pas en  $x_0$ , et que  $f''_{\varepsilon}(x_0; d, d) \leq 0 \forall d \in \mathbb{R}^n$

D'abord 
$$f''_{\varepsilon}(x_0; d, d) = 0$$

en effet: 
$$f''_{\varepsilon}(x_0; d, d) = \min_{\eta \in M_{d, \varepsilon}(x_0)} \left\{ \eta [f'_{\varepsilon}(x_0; d) - f'(x_0; d)] \right\}$$

OR  $f'_{\varepsilon}(x_0; d) - f'(x_0; d) > 0$

et  $0 \in M_{d, \varepsilon}(x_0)$ .

Dès lors  $f''_{\varepsilon}(x_0; d, d) = 0$  □

Ensuite 
$$f''_{\varepsilon}(x_0; d, d) \leq 0 \quad \forall d \in \mathbb{R}^n$$

en effet: 
$$f''_{\varepsilon}(x_0; d, d) = \min_{\eta \in M_{d, \varepsilon}(x_0)} \max_{x^* \in \partial f(x_0)} \left\{ \eta [\langle x^*, d \rangle - f'(x_0; d)] \right\}$$

OR  $\forall x^* \in \partial f(x_0) \quad \langle x^*, d \rangle \leq f'_{\varepsilon}(x_0; d)$

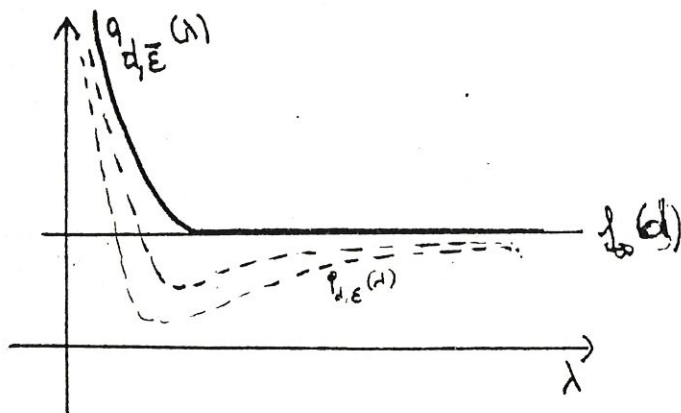
D'où  $f''_{\varepsilon}(x_0; d, d) \leq \min_{\eta \in M_{d, \varepsilon}(x_0)} \eta [f'_{\varepsilon}(x_0; d) - f'(x_0; d)]$

où le second membre de l'inégalité est nul. □

Il ne peut que  $\langle x^*, \delta \rangle < f'(x_0; \delta)$  pour un certain  $x^* \in \partial_\varepsilon f(x_0)_d$ . On peut dès lors assurer qu'il existe une direction  $\delta \in \mathbb{R}^n$  telle que

$$f''_\varepsilon(x_0; d, \delta) < 0$$

Remarquons que, si  $d$  satisfait  $(S_2)$  en  $x_0$  pour un certain  $\bar{\varepsilon}$ , i.e. si  $q_{d, \bar{\varepsilon}}(\lambda)$  "s'arrête" sur la droite  $f_\infty(d)$ , alors  $d$  satisfait  $(S_1)$  pour tout  $\varepsilon < \bar{\varepsilon}$  puisque la fonction  $q_{d, \varepsilon}(\lambda)$  est décroissante en  $\varepsilon$ .

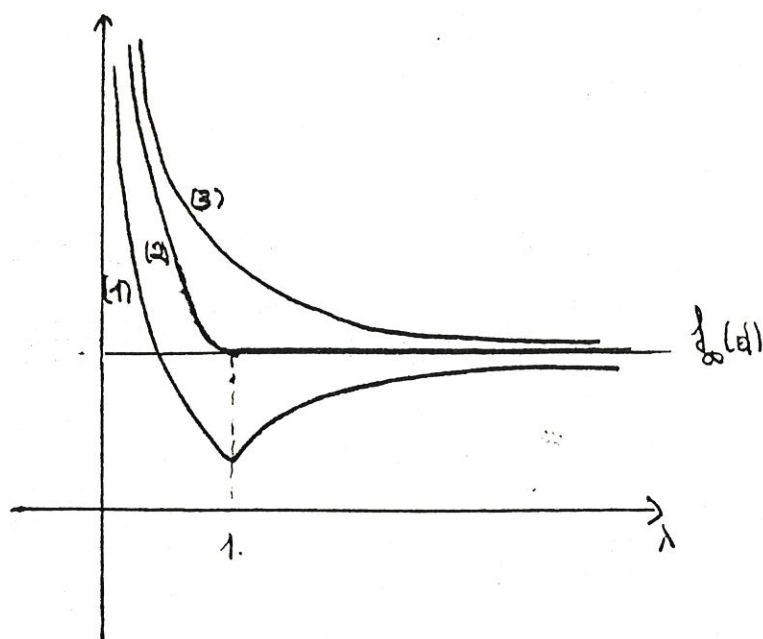


L'exemple 4, que nous rappelons ci-dessous, nous suggère un tel comportement de  $q_{d, \varepsilon}(\lambda)$ .

$f: \mathbb{R} \longrightarrow (-\infty, +\infty]$  définie par

$$f(x) = \begin{cases} x & \text{si } 0 \leq x \leq 1 \\ 2x-1 & \text{si } x > 1 \\ +\infty & \text{si } x < 0 \end{cases}$$

Suivant que  $\varepsilon$  est  $< 1$ ,  $= 1$ , ou  $> 1$ , nous avons les trois situations distinctes:



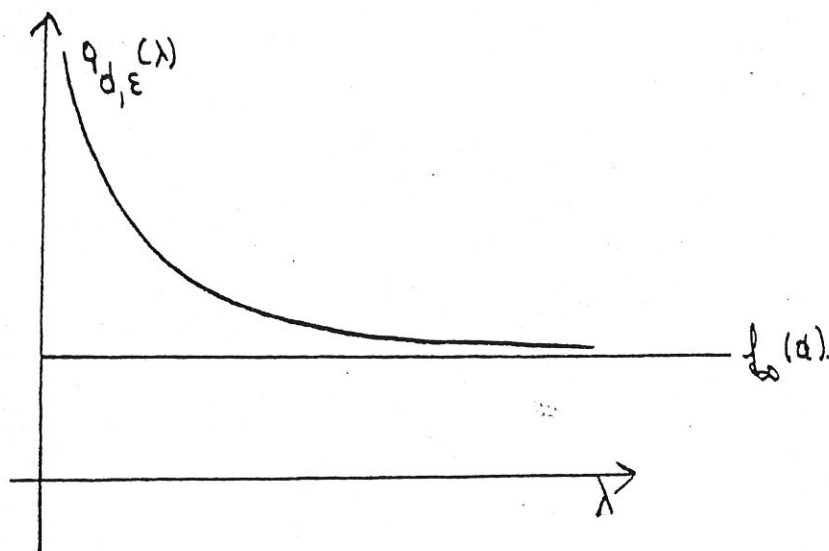
Les courbes (1) et (2) reflètent des situations déjà étudiées ( $S_1$ ) et ( $S_2$ ). Nous avons complété ce graphe en y ajoutant la courbe (3) qui représente  $q_{d,\varepsilon}(1)$  pour  $\varepsilon > 1$ . Notons que'elle se situe toujours au dessus de la droite  $f_\infty(d)$ . L'étude théorique de ce dernier cas complètera l'éventail des différentes caractéristiques de  $\Lambda_{d,\varepsilon}(x_0)$ .

Troisième cas :  $\Lambda_{d,\varepsilon}(x_0)$  est vide.

Le cas où  $\Lambda_{d,\varepsilon}(x_0)$  est vide correspond à la situation où pour tout  $\lambda > 0$

$$\boxed{S_3} \quad \frac{f(x_0 + \lambda d) - f(x_0) + \varepsilon}{\lambda} > f_\infty(d)$$

Cette situation, notée ( $S_3$ ), se représente comme suit :



L'ensemble  $M_{d,\epsilon}(x_0)$  est alors réduit au singleton  $\{0\}$ .  
Le théorème 3.1 nous permet donc d'assurer que  $v_{d,\epsilon}$  est différentiable en  $x_0$  et que son gradient est nul.

Notons que dans ce cas, nous avons  $f'_\epsilon(x_0; d) = f_{\infty}(d)$ .

On pourrait s'interroger sur la validité éventuelle de l'égalité  $f'_\epsilon(x_0; d) = f'(x_0; d)$ .

Par le lemme 3.2, nous savons que, lorsque  $\Lambda_{d,\epsilon}(x_0)$  est non vide,  $f'_\epsilon(x_0; d) > f'(x_0; d)$ . Dans la situation présente, où  $\Lambda_{d,\epsilon}(x_0)$  est vide, on ne peut être si catégorique.

En effet, l'inégalité stricte ou l'égalité peuvent se présenter. Nous pouvons donc séparer  $(S_3)$  en 2 sous-cas:

$(S_{3a})$  et  $(S_{3b})$ .

—  $(S_{3a})$  :  $f'_\epsilon(x_0; d) > f'(x_0; d)$

$(S_{3a})$  est un état "transitoire", c-à-d quand  $\epsilon$  décroît



vers 0, on passe d'abord par  $(S_2)$  et ensuite par  $(S_1)$ .  
Ceci était suggéré par l'exemple 4 rappelé ci-dessus.

En effet, nous avons :

$$f_\infty(d) = 2$$

$$f'_\varepsilon(x_0; d) = \begin{cases} 2 & \text{si } \varepsilon \geq 1 \\ 1 + \varepsilon & \text{si } 0 < \varepsilon < 1 \end{cases}$$

$$f'(x_0; d) = 1.$$

D'une part  $\forall \varepsilon > 0 \quad f'_\varepsilon(x_0; d) > f'(x_0; d)$

D'autre part, quand  $\varepsilon$  décroît, nous passons successivement par des situations  $(S_3)$  ( $\cap_{d,\varepsilon}(x_0) = \emptyset$ ),  $(S_2)$  ( $\cap_{d,\varepsilon}(x_0) = ]1, \infty[$ ) et  $(S_1)$  ( $\cap_{d,\varepsilon}(x_0) = \{1\}$ ).

$$- (S_{3b}) : \underline{f'_\varepsilon(x_0; d) = f'(x_0; d) = f_\infty(d)}.$$

$$\text{Comme } f'(x_0; d) = \lim_{\lambda \rightarrow 0} \frac{f(x_0 + \lambda d) - f(x_0)}{\lambda} \quad \text{et}$$

$$f_\infty(d) = \lim_{\lambda \rightarrow \infty} \frac{f(x_0 + \lambda d) - f(x_0)}{\lambda}$$

nous avons  $q_d(\lambda) = f'(x_0; d) \quad \forall \lambda$  ou encore

$$f(x_0 + \lambda d) = f(x_0) + \lambda f'(x_0; d).$$

$$\text{La fonction } f_d : \begin{matrix} \mathbb{R}_0^+ & \longrightarrow & \mathbb{R} \\ \lambda & \longrightarrow & f(x_0 + \lambda d). \end{matrix}$$

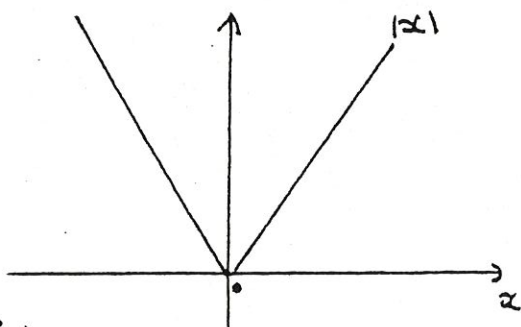
est donc affine sur  $\mathbb{R}_0^+$ .

On peut dès lors reformuler  $q_{d,\varepsilon}(\lambda)$  comme  $f'(x_0; d) + \frac{\varepsilon}{\lambda}$

Grâce à cette expression, on peut remarquer que

la situation  $(S_{3b})$  est, contrairement à  $(S_{3a})$ , un état permanent car  $q_{d,\varepsilon}(\lambda)$  a le même comportement pour tout  $\varepsilon$ .

Explicitons ce dernier cas par l'exemple (8) de la fonction  $|x|$  de  $\mathbb{R}$  dans  $\mathbb{R}$ .



Choisissons  $x_0 = 0$

$$q_{d,\varepsilon}(\lambda) = \frac{|\lambda d| + \varepsilon}{\lambda} = |d| + \frac{\varepsilon}{\lambda}$$

$$\text{Donc. } f_{\infty}(d) = |d|.$$

$$\cdot f'_{\varepsilon}(x_0; d) = \inf_{\lambda > 0} q_{d,\varepsilon}(\lambda) = |d| = f_{\infty}(d)$$

$$q_d(\lambda) = |d| = f'_{\varepsilon}(x_0; d).$$

Ceci reflète l'état permanent de la situation  $(S_{3b})$

Ce dernier cas complète la présentation des différentes situations possibles concernant l'ensemble  $\Lambda_{d,\varepsilon}(x_0)$ :  $(S_1)$  quand il est non vide et borné,  $(S_2)$  quand il est non vide et non borné,  $(S_3)$  quand il est vide.

Dans ce chapitre, nous nous sommes proposé de définir une dérivée directionnelle seconde d'une fonction  $f$  convexe, non nécessairement différentiable en  $x_0$ :  $f''_{\varepsilon}(x_0; d, \delta)$ . Le premier paragraphe nous a fourni un théorème fondamental sur son expression.

Comme  $f''_{\varepsilon}(x_0; d, \delta) = \langle \nabla v_{d, \varepsilon}(x_0), \delta \rangle$  lorsque la fonction  $v_{d, \varepsilon}(\cdot)$  est différentiable au point  $x_0$ , nous avons énoncé, au second paragraphe des hypothèses assurant l'existence de son gradient. Dans le troisième paragraphe, nous avons scindé l'étude de l'ensemble  $\Pi_{d, \varepsilon}(x_0)$  en trois situations. Chacune d'elles nous a fourni des résultats plus précis sur l'expression du gradient de  $v_{d, \varepsilon}$ . Elles nous permettent d'émettre des éléments de réponse à la question principale de ce mémoire: "Dans quelle mesure cette dérivée directionnelle seconde  $f''_{\varepsilon}(x_0; d, \delta)$  est-elle une approximation de la dérivée seconde au sens habituel de  $f$ ? Grâce aux notions développées jusqu'ici, nous pouvons aborder ce problème dans le chapitre suivant.

## CHAPITRE IV

Comportement à la limite de  
 $f''_{\varepsilon}(x_0; d, \delta)$  quand  $\varepsilon \rightarrow 0$ .



Au cours du chapitre précédent, nous avons désigné par  $v_{d,\varepsilon}(\cdot)$  l'application qui à  $x$  associe  $f'_\varepsilon(x; d)$ . En un point  $x_0$  où  $v_{d,\varepsilon}(\cdot)$  n'est pas différentiable, elle admet néanmoins une dérivée directionnelle dans toute direction  $\delta$ , que nous avons notée par  $f''_\varepsilon(x_0; d, \delta)$ . Rappelons que

$$f''_\varepsilon(x_0; d, \delta) = \min_{\mu \in M_{d,\varepsilon}(x_0)} \left\{ \mu \left[ \psi_{\frac{\delta}{\|\delta\|}}^* f(x_0) - f'(x_0; \delta) \right] \right\}$$

où  $\psi_A^*$  est la fonction d'appui de  $A$ .

Dans le cas diagonal,  $f''_\varepsilon(x_0; d, d) = \bar{\mu}_{d,\varepsilon}(x_0) [f'_\varepsilon(x_0; d) - f'(x_0; d)]$

où  $\bar{\mu}_{d,\varepsilon}(x_0)$  est le minimum de l'ensemble  $M_{d,\varepsilon}(x_0)$ .

Le comportement de  $f''_\varepsilon(x_0; d, \delta)$  comme fonction de  $\varepsilon$  va donc dépendre des propriétés de  $f'_\varepsilon(x_0; d)$  et de  $M_{d,\varepsilon}(x_0)$  comme application et multiapplication de  $\varepsilon$ . Dès lors, nous noterons dans la suite  $M_{d,\varepsilon}(x_0)$  par  $M_d(\varepsilon)$  et  $\bar{\mu}_{d,\varepsilon}(x_0)$  par  $\bar{\mu}_d(\varepsilon)$ .

Nous sommes donc amenés dans un premier paragraphe, à traiter du comportement de  $M_d(\varepsilon)$  et de  $\bar{\mu}_d(\varepsilon)$  quand  $\varepsilon$  tend vers zéro.

Nous précisons les trois situations déjà mises en évidence:  $\bar{\mu}_d(\varepsilon)$  peut être non vide et borné, non vide et non borné ou vide. Dans un second paragraphe, nous examinerons le comportement de  $f'_\varepsilon(x_0; d)$  quand  $\varepsilon$  tend vers l'infini et vers zéro. La troisième partie tentera de montrer dans quel sens  $f''_\varepsilon(x_0; d, \delta)$  approxime la dérivée seconde de  $f$  en  $x_0$ .

#### IV. 1 Etude de $\Lambda_d(\varepsilon)$ en fonction de $\varepsilon$

---

##### A. Nouvelle description de $\Lambda_d(\varepsilon)$

---

Nous savons que l'on désigne par  $\Lambda_d(\varepsilon)$  l'ensemble (éventuellement vide) des  $\lambda_0 > 0$  pour lesquels le quotient différentiel approche

$$q_{d,\varepsilon}(\lambda) : \lambda \rightarrow \frac{f(x_0 + \lambda d) - f(x_0) + \varepsilon}{\lambda}$$

atteint son minimum sur  $\mathbb{R}_+^*$ .

Comme  $q_d(\lambda)$  est une fonction croissante en  $\lambda$ ,  $\Lambda_d(0)$  n'est autre que  $]0, a_d^*]$  où  $a_d^* \in \overline{\mathbb{R}}_+$  est le supremum de tous les  $a \in \mathbb{R}$  tels que

$$\frac{f(x_0 + \lambda d) - f(x_0)}{\lambda} = f'(x_0; d) \quad \forall \lambda \in ]0, a]$$

ou encore

$$f(x_0 + \lambda d) = f(x_0) + \lambda f'(x_0; d) \quad \forall \lambda \in [0, a] \quad (4.1.1)$$

##### A.1. Interprétation de $a_d^*$

---

1)  $a_d^* = +\infty$

Ceci correspond à la situation où  $f_d : \lambda \rightarrow f(x_0 + \lambda d)$  est affine sur  $\mathbb{R}^+$ . En effet, (4.1.1) nous fournit

$$f(x_0 + \lambda d) = f(x_0) + \lambda f'(x_0; d) \quad \forall \lambda$$

C'est le cas de la fonction  $|x|$  de  $\mathbb{R}$  dans  $\mathbb{R}$ . Prenons  $x_0 = 0$  et  $d = 1$ . Alors nous avons

$$f(x_0 + \lambda d) = \lambda = f(x_0) + \lambda f'(x_0; d)$$

et donc  $q_d(\lambda) = f'(x_0; d) \quad \forall \lambda$  i.e.  $f$  est affine sur la droite  $x_0 + \mathbb{R}_0^+ d$ .

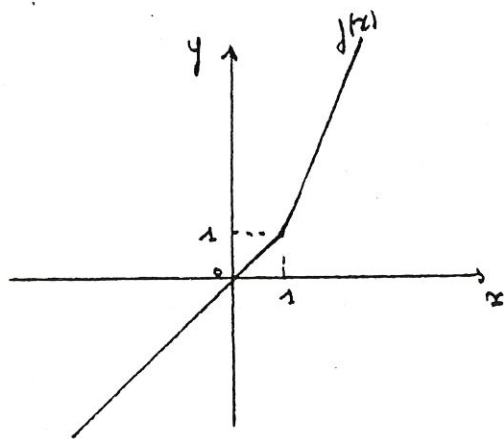
2)  $0 < a_d^* < +\infty$

Dans ce cas  $f_d(\lambda)$  est affine sur le segment  $[0, a_d^*]$  ou encore  $f$  est affine sur  $[x_0, x_0 + a_d^* d]$

exemple 3

Soit  $f: \mathbb{R} \rightarrow \mathbb{R}$

$$f(x) = \begin{cases} x & \text{si } x \leq 1 \\ 2x - 1 & \text{si } x \geq 1 \end{cases}$$



Choisissons  $x_0 = 0$ ,  $d = 1$ , calculons  $q_d(\lambda)$  et  $f'(x_0; d)$ .

$$q_d(\lambda) = \begin{cases} 1 & \text{si } \lambda \leq 1 \\ \frac{2\lambda - 1}{\lambda} & \text{si } \lambda \geq 1 \end{cases}$$

$$f'(x_0; d) = \lim_{\lambda \rightarrow 0} q_d(\lambda) = 1$$

On voit que  $\forall \lambda \leq 1 \quad f(x_0 + \lambda d) = f(x_0) + \lambda f'(x_0; d)$   
 et donc  $a_d^* = 1$ . On pourrait s'y attendre à la vue du  
 graphe, puisque  $f$  est affine sur  $[0, 1]$ .

3)  $a_d^* = 0$

Ce cas où  $f$  n'est pas affine sur la demi-droite  
 $x_0 + \mathbb{R}^+ d$  est équivalent à l'hypothèse H :

$$(H) \quad \exists a > 0 \text{ t.q. } q_d(\lambda) > f'(x_0; d) \quad \forall \lambda \in ]0, a]$$

Reprenons l'exemple 5 déjà traité

$$f: \mathbb{R} \rightarrow \mathbb{R}$$

$$x \mapsto \begin{cases} x^2 & \text{si } x \leq 0 \\ x^2 + x & \text{si } x \geq 0 \end{cases} \quad \text{où } x_0 = 0 \text{ et } d = 1$$

Nous avons déjà calculé que

$$q_d(\lambda) = \lambda + 1 \quad \text{et} \quad f'(0; 1) = 1$$

Il n'existe donc pas de  $\lambda > 0$  t.q.  $q_d(\lambda) = f'(x_0; d)$

De fait,  $f$  n'est pas affine sur  $\mathbb{R}^+$ .

La valeur  $a_d^*$  reflète donc le caractère affine de la fonction  $f$   
 sur le segment  $[x_0, x_0 + a_d^* d]$ . En résumé:

- si  $a_d^* = +\infty$ ,  $f$  est affine sur  $x_0 + \mathbb{R}^+ d$
- si  $0 < a_d^* < +\infty$ ,  $f$  est affine sur  $[x_0, x_0 + a_d^* d]$
- si  $a_d^* = 0$ ,  $f$  n'est pas affine sur  $x_0 + \mathbb{R}^+ d$



## A.2 Conséquences sur la classification de $\Lambda_d(\varepsilon)$ :

Les différentes propriétés de  $\Lambda_d(\varepsilon)$ , ou de manière équivalente celles de  $M_d(\varepsilon)$  dépendent de  $q_{d,\varepsilon}(R_0^+)$  ainsi que de  $f_\infty(d)$ , en gardant à l'esprit que :  $\Lambda_d(\varepsilon) = \left\{ \frac{1}{\mu_0} \mid \mu_0 \in M_d(\varepsilon), \mu_0 > 0 \right\}$ . La classification de tous les cas possibles a été faite au chapitre précédent, en conjonction avec les propriétés de différentiabilité de la fonction  $x \rightarrow f'_\varepsilon(x; d)$ . Elle est améliorée par l'introduction de  $a_d^*$ .

1)  $a_d^* = +\infty$

Cette situation peut être décrite par l'un de énoncés équivalents suivants :

i)  $f'_\varepsilon(x_0; d) = f'(x_0; d) \quad \forall \varepsilon > 0$

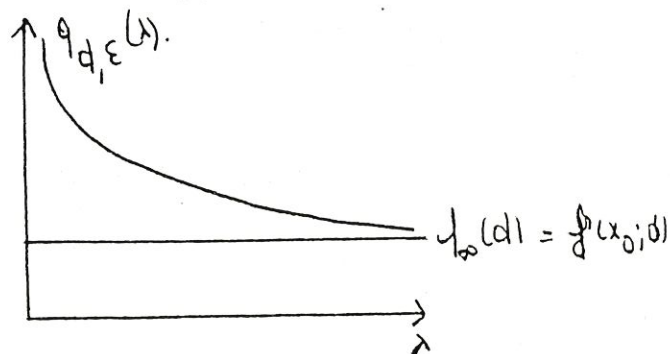
ii)  $\Lambda_d(\varepsilon) = \emptyset \quad \forall \varepsilon > 0$

iii)  $M_d(\varepsilon) = \{0\} \quad \forall \varepsilon > 0$

iv)  $\left\{ q_{d,\varepsilon}(\lambda) \mid \lambda > 0 \right\} \cap ]-\infty, f_\infty(d)] = \emptyset \quad \forall \varepsilon > 0$

v)  $\Lambda_d(0) = R_0^+$

Ceci nous rappelle la situation  $S_{3b}$  du chapitre III.(3)



2)  $\underline{a_d^*} < +\infty$

Dans ce cas il existe  $\bar{\varepsilon} > 0$  tel que

$$q_{d, \bar{\varepsilon}}(R_0^+) \cap ]-\infty, f_\infty(d)] \neq \emptyset \quad (4.1.2).$$

On définit ensuite  $\varepsilon_d^* > 0$  comme le supremum de tous les  $\bar{\varepsilon} > 0$  pour lesquels la relation (4.1.2) est satisfaite.

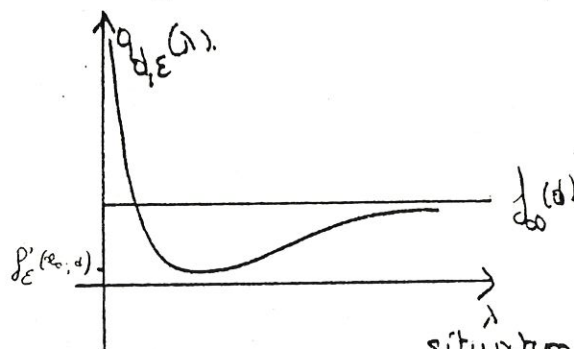
La situation peut être alors complètement décrite comme suit:

— (a) si  $\varepsilon \in ]0, \varepsilon_d^* [$  alors on a les équivalences

i)  $f'_\varepsilon(x_0; d) < f'_\varepsilon(x_0; d) < f_\infty(d)$

ii)  $M_d(\varepsilon)$  est un intervalle compact non vide

iii)  $0 \notin M_d(\varepsilon)$



situation  $S_1$  de III.3

Remarquons que l'on ne peut avoir l'égalité

$$f'_\varepsilon(x_0; d) = f_\infty(d) \text{ car alors pour } \varepsilon < \varepsilon' < \varepsilon_d^* \text{ on}$$

$$\text{aura } \{ q_{d, \varepsilon'}(\lambda) \mid \lambda \in R_0^+ \} \cap ]-\infty, f_\infty(d)] = \emptyset,$$

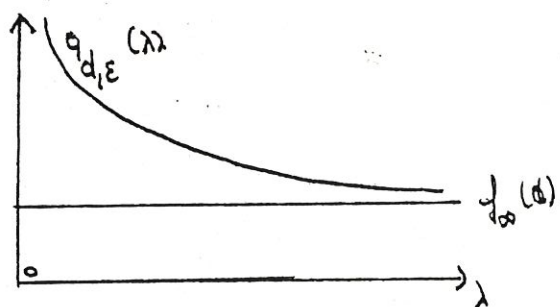
ce qui contredit la définition de  $\varepsilon_d^*$ .

— (b) si  $\varepsilon \in ]\varepsilon_d^*, +\infty[$  alors on a les équivalences

i)  $f'(x_0; d) < f'_\varepsilon(x_0; d) = f'_\infty(d)$

ii)  $\Lambda_d(\varepsilon)$  est vide

iii)  $M_d(\varepsilon) = \{0\}$



situation  $S_{3a}$  de III.3

— (c) si  $\varepsilon = \varepsilon_d^*$  alors on a les équivalences

i)  $f'(x_0; d) < f'_{\varepsilon_d^*}(x_0; d) = f'_\infty(d)$

ii)  $\Lambda_d(\varepsilon_d^*)$ , s'il est non vide, est non borné

iii)  $M_d(\varepsilon_d^*)$  contient 0

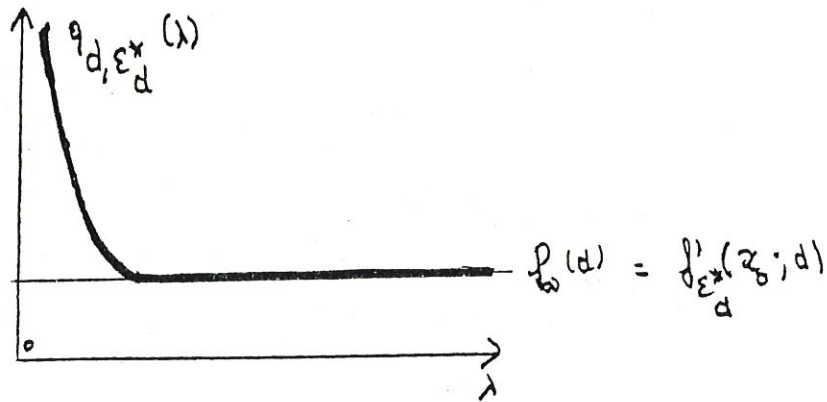
Remarquons que  $\Lambda_d(\varepsilon_d^*)$  est vide quand  $\varepsilon_d^*$ , étant défini comme un supremum, ne vérifie pas la relation (4.1.2).  $\Lambda_d(\varepsilon_d^*)$  est non vide et non borné quand il existe un  $\bar{\lambda}$  tel que.

$$q_{d, \varepsilon_d^*}(\lambda) = f'_\infty(d) \quad \forall \lambda \geq \bar{\lambda}$$

c.a'-d. tel que

$$f_d(\lambda) = f(x_0 + \lambda d) = f(x_0) + \lambda f'_\infty(d) - \varepsilon \quad \forall \lambda \geq \bar{\lambda}.$$

Donc  $\Lambda_d(\varepsilon_d^*)$  est non vide, et donc non borné, si  $f_d(\cdot)$  est affine pour  $\lambda$  suffisamment grand.



### Remarque

Il est commode de poser  $\varepsilon_d^* = 0$  pour le premier cas envisagé (i.e.  $f_d$  affine sur  $\mathbb{R}_0^+$ ). De plus, il est aisé de voir que  $\varepsilon_d^* < +\infty$  si et seulement si  $f_{\infty}(d) < +\infty$ .

En effet,  $f_{\infty}(d) = +\infty$  si la relation (4.12) a lieu  $\forall \varepsilon$  et donc  $\varepsilon_d^* = +\infty$  (4.1.3).

Pour concrétiser cette classification, reprenons l'exemple où  $f$  est définie comme suit:

$$f: \mathbb{R} \longrightarrow \bar{\mathbb{R}}$$

$$x \longmapsto \begin{cases} x & \text{si } 0 \leq x \leq 1 \\ 2x-1 & \text{si } x > 1 \\ +\infty & \text{si } x < 0 \end{cases}$$

Choisissons  $x_0 = 0$ ,  $d = 1$ .

Le calcul de  $a_d^*$  nous permettra de nous situer dans la classification. Comme

$$a_d^* = 1 < +\infty, \text{ nous pouvons donc calculer le } \varepsilon_d^*$$



Rappelons que  $f_\infty(d) = 2$  et

$$q_{d,\varepsilon}(\lambda) = \begin{cases} 1 + \frac{\varepsilon}{\lambda} & \text{si } 0 \leq \lambda \leq 1 \\ 2 + \frac{\varepsilon-1}{\lambda} & \text{si } \lambda \geq 1 \end{cases}$$

Si  $\varepsilon > 1$  alors  $\{q_{d,\varepsilon}(\lambda) \mid \lambda \in \mathbb{R}_0^+\} \cap ]-\infty, 2] = \emptyset$

Si  $\varepsilon \leq 1$  alors  $\{q_{d,\varepsilon}(\lambda) \mid \lambda \in \mathbb{R}_0^+\} \cap ]-\infty, 2] \neq \emptyset$

Dès lors  $\varepsilon_d^* = 1$

— (a) Si  $\varepsilon \in ]0, 1[$  alors

$$\begin{aligned} \text{i) } f'_d(x_0; d) &< f'_\varepsilon(x_0; d) < f_\infty(d) \\ 1 &< 1 + \varepsilon < 2 \end{aligned}$$

ii)  $\Lambda_d(\varepsilon) = \{1\}$  est un intervalle compact non vide  
et  $M_d(\varepsilon) = \{1\}$

— (b) si  $\varepsilon \in ]1, +\infty[$  alors

$$\begin{aligned} \text{i) } f'_d(x_0; d) &< f'_\varepsilon(x_0; d) = f_\infty(d) \\ 1 &< 2 = 2 \end{aligned}$$

ii)  $\Lambda_d(\varepsilon) = \emptyset$  et  $M_d(\varepsilon) = \{0\}$

— (c) Si  $\varepsilon = 1$  alors

$$\text{i) } f'_\varepsilon(x_0; d) = f_\infty(d)$$

ii)  $\Lambda_d(\varepsilon) = [1, +\infty[$  est non vide et non borné

iii)  $M_d(\varepsilon) = [0, 1]$  contient 0.

### B. Comportement de $\Lambda_d(\varepsilon)$ quand $\varepsilon$ tend vers 0

A première vue, on pourrait espérer que  $\bar{\lambda}_\varepsilon$  ou  $\underline{\lambda}_\varepsilon$  (qui sont respectivement le maximum et le minimum de  $\Lambda_d(\varepsilon)$ ) converge vers 0 quand  $\varepsilon$  tend vers 0<sup>+</sup>. En effet, quand  $\varepsilon = 0$ , le  $\lambda$  qui minimise  $q_d(\lambda)$  est 0 puisque  $q_d(\cdot)$  est croissante en  $\lambda$ . Mais ceci n'est pas vrai en général. Il suffit, pour s'en convaincre, de considérer l'exemple précédent. En effet,  $\forall \varepsilon \in ]0, 1[$   $\Lambda_d(\varepsilon) = \{1\}$  et donc  $\lim_{\varepsilon \rightarrow 0} \Lambda_d(\varepsilon) = \{1\}$ . Remarquons que cette limite est précisément la valeur de  $a_d^*$ . Ceci est une amorce au résultat du théorème qui sera prouvé grâce aux deux propositions suivantes.

#### Proposition 4.1.1.

Soient  $\underline{\lambda}_\varepsilon$  et  $\bar{\lambda}_\varepsilon$  les minimum et maximum de

$\Lambda_d(\varepsilon)$  pour tout  $\varepsilon < \bar{\varepsilon}$

$$\text{Alors a) } f'_\varepsilon(x_0; d) - f'(x_0; d) \geq \frac{\varepsilon}{\underline{\lambda}_\varepsilon} \quad (4.1.4)$$

de telle sorte que

$$\lim_{\varepsilon \rightarrow 0} \frac{\varepsilon}{\underline{\lambda}_\varepsilon} = 0$$

$$\text{b) } \frac{1}{\bar{\lambda}_{\varepsilon'}} < \frac{1}{\underline{\lambda}_\varepsilon} \quad \text{pour tout } \varepsilon' < \varepsilon \quad (4.1.5)$$

### Preuve

a) D'après la définition de  $f'_\varepsilon(x_0; d)$ , il résulte que

$$\frac{f(x_0 + \bar{\lambda}_{\varepsilon'} d) - f(x_0) + \varepsilon'}{\bar{\lambda}_{\varepsilon'}} \leq \frac{f(x_0 + \underline{\lambda}_\varepsilon d) - f(x_0) + \varepsilon}{\underline{\lambda}_\varepsilon} \quad \text{car } \bar{\lambda}_{\varepsilon'} \in \eta_d(\varepsilon')$$

$$\geq \frac{f(x_0 + \underline{\lambda}_\varepsilon d) - f(x_0) + \varepsilon}{\underline{\lambda}_\varepsilon} + \frac{\varepsilon' - \varepsilon}{\underline{\lambda}_\varepsilon}$$

D'où

$$f'_\varepsilon(x_0; d) \leq f'_\varepsilon(x_0; d) + \frac{\varepsilon' - \varepsilon}{\underline{\lambda}_\varepsilon}$$

De manière similaire, on obtient en échangeant  $\varepsilon$  et  $\varepsilon'$ :

$$f'_\varepsilon(x_0; d) \leq f'_{\varepsilon'}(x_0; d) + \frac{\varepsilon - \varepsilon'}{\bar{\lambda}_{\varepsilon'}}$$

En combinant ces deux inégalités, on obtient :

$$\frac{\varepsilon - \varepsilon'}{\underline{\lambda}_\varepsilon} \leq f'_\varepsilon(x_0; d) - f'_{\varepsilon'}(x_0; d) \leq \frac{\varepsilon - \varepsilon'}{\bar{\lambda}_{\varepsilon'}} \quad (4.1.6)$$

En faisant tendre  $\varepsilon'$  vers 0 dans la première inégalité de (4.1.6), nous tirons :

$$f'_\varepsilon(x_0; d) - f'_\varepsilon(x_0; d) \geq \frac{\varepsilon}{\underline{\lambda}_\varepsilon}$$

b) De (4.1.6) il découle

$$\frac{\varepsilon - \varepsilon'}{\underline{\lambda}_\varepsilon} \leq \frac{\varepsilon - \varepsilon'}{\bar{\lambda}_{\varepsilon'}} \quad \text{et donc}$$

$$\forall \varepsilon' < \varepsilon \quad \text{on a } \bar{\lambda}_{\varepsilon'} < \underline{\lambda}_\varepsilon \quad \square$$

Les résultats de cette proposition suggèrent deux remarques.

### Remarque 1

Cette proposition nous assure que des ensembles  $\Lambda_d(\varepsilon)$  ont au plus un point commun.

En effet, prenons  $\lambda_\varepsilon$  et  $\lambda_{\varepsilon'}$  quelconques dans  $\Lambda_d(\varepsilon)$  et  $\Lambda_d(\varepsilon')$ . Utilisant respectivement la définition de  $\overline{\lambda}_\varepsilon$ , la proposition (4.1.1) et la définition de  $\underline{\lambda}_{\varepsilon'}$ , on obtient

$$\lambda_\varepsilon \leq \overline{\lambda}_\varepsilon \leq \underline{\lambda}_{\varepsilon'} \leq \lambda_{\varepsilon'} \quad (4.1.7)$$

et donc  $\Lambda_d(\varepsilon)$  et  $\Lambda_d(\varepsilon')$  sont disjoints ou n'ont qu'un seul point en commun.

### Remarque 2

Si  $\underline{\lambda}_\varepsilon$  converge vers 0, alors il converge moins vite que  $\varepsilon$ , quand celui-ci tend vers 0.

Ainsi, la différence  $f'_\varepsilon(x_0; d) - f'(x_0; d)$  peut tendre lentement vers 0. Le cas de la fonction quadratique (exemple 6) nous le montre: cette différence vaut  $\sqrt{2\varepsilon \langle Ad, d \rangle}$ . D'ordre  $\varepsilon^{1/2}$ , elle décroît moins vite que  $\varepsilon$ . De même, l'unique  $\lambda_\varepsilon$  qui vaut  $\sqrt{\frac{2\varepsilon}{\langle Ad, d \rangle}}$ , converge plus lentement vers 0 que  $\varepsilon$ .

Enonçons la deuxième proposition, dont la démonstration se trouve en ([2], p. 21).



### Proposition 4.1.2

Soit  $\{\varepsilon_m\}_m$  une suite convergente vers  $\varepsilon \in ]0, \bar{\varepsilon}[$

Alors une des assertions suivantes est vérifiée.

$$a) \lim_{m \rightarrow \infty} \underline{\lambda}_{\varepsilon_m} = \lim_{m \rightarrow \infty} \bar{\lambda}_{\varepsilon_m} = \underline{\lambda}_{\varepsilon}$$

$$b) \lim_{m \rightarrow \infty} \underline{\lambda}_{\varepsilon_m} = \lim_{m \rightarrow \infty} \bar{\lambda}_{\varepsilon_m} = \bar{\lambda}_{\varepsilon}$$

$$c) \left[ \liminf_{m \rightarrow \infty} \underline{\lambda}_{\varepsilon_m}, \limsup_{m \rightarrow \infty} \bar{\lambda}_{\varepsilon_m} \right] = \Lambda_d(\varepsilon)$$

Nous sommes maintenant en mesure de démontrer le résultat principal de convergence de  $\Lambda_d(\varepsilon)$  quand  $\varepsilon$  tend vers 0.

### Théorème 4.1

Si  $a_d^* < +\infty$  en  $x_0$  pour la direction  $d$

Alors  $\Lambda_d(\varepsilon)$  est non vide pour  $\varepsilon$  dans un voisinage de 0 et

$$\lim_{\varepsilon \rightarrow 0} \Lambda_d(\varepsilon) = \{a_d^*\} \quad (4.1.8)$$

où (4.1.8) doit être interprété comme

$$\lim_{\varepsilon \rightarrow 0} \underline{\lambda}_{\varepsilon} = \lim_{\varepsilon \rightarrow 0} \bar{\lambda}_{\varepsilon} = a_d^*$$

### Preuve

Soit  $a_d^* < +\infty$  en  $x_0$  pour la direction  $d$ .

Rappelons que  $a_d^*$  est le supremum de tous les  $a \in \mathbb{R}^+$  pour lesquels  $f(x_0 + \lambda d) = f(x_0) + \lambda f'(x_0; d) \quad \forall \lambda \in [0, a]$

Par hypothèse, nous écartons la situation dégénérée où  $\lambda_d(\varepsilon)$  est vide pour tout  $\varepsilon$ .

Dès lors  $\lambda_d(\varepsilon)$  est non vide pour  $\varepsilon$  dans un voisinage de  $0^+$  noté  $]0, \varepsilon]$ .

Soit une suite  $\{\varepsilon_m\}_m$  convergente vers  $0^+$ . Celle-ci génère une suite  $\{\lambda_{\varepsilon_m}\}$  telle que  $\lambda_{\varepsilon_m} \in \lambda_d(\varepsilon_m) \quad \forall m$  ou encore telle que  $f(x_0 + \lambda_{\varepsilon_m} d) = f(x_0) + \lambda_{\varepsilon_m} f'_{\varepsilon_m}(x_0; d) - \varepsilon_m \quad \forall m \quad (4.1.9)$

Par la proposition (4.1.1), la suite  $\{\lambda_{\varepsilon_m}\}_m$  est bornée.

On peut donc en extraire une sous-suite convergente, encore notée  $\{\lambda_{\varepsilon_m}\}_m$ . Désignons sa limite par  $\lambda_\infty$ .

Prouvons maintenant par la méthode de la double inégalité que  $\lambda_\infty = a_d^*$ . La valeur d'adhérence  $\lambda_\infty$  étant quelconque, la proposition 4.1.2 nous permettra de conclure que  $\lim_{\varepsilon \rightarrow 0} \lambda_d(\varepsilon) = \{a_d^*\}$

1) Puisque  $f_d(\lambda_{\varepsilon_m})$  est continue sur  $\mathbb{R}_0^+$  et que  $f'_{\varepsilon_m}(x_0; d)$  tend vers  $f'(x_0; d)$  on obtient, en passant à la limite dans (4.1.9) :

$$f(x_0 + \lambda_\infty d) = f(x_0) + \lambda_\infty f'(x_0; d)$$

et dès lors  $\lambda_\infty \leq a_d^*$ .

En particulier, dans le cas où  $a_d^* = 0$ , nous avons directement que  $\lambda_\infty = 0$  et donc d'égalité de deux termes.

2) La fonction convexe  $\pi_{d,\varepsilon} : \mathcal{Y} \longrightarrow \mathcal{Y} \cdot [f(x_0 + \frac{d}{\mathcal{Y}}) - f(x_0) + \varepsilon]$  est telle que

$$\pi_{d,\varepsilon}(\mathcal{Y}) = f'(x_0; d) + \varepsilon \mathcal{Y} \quad \forall \mathcal{Y} \in M_d(0)$$

$$\text{où } M_d(0) = \left[ \frac{1}{a_d^*}, +\infty \right[$$

$$\text{D'autre part } M_d(\varepsilon) \subset \left] 0, \frac{1}{a_d^*} \right] \quad \forall \varepsilon \in ]0, \bar{\varepsilon}[$$

en effet soit  $\mathcal{Y} \in M_d(\varepsilon)$ .

$$\text{alors } \pi_{d,\varepsilon}(\mathcal{Y}) = f'_\varepsilon(x_0; d)$$

La convexité de la fonction  $\pi_{d,\varepsilon}(\cdot)$  et sa croissance stricte sur  $\left[ \frac{1}{a_d^*}, +\infty \right[$  nous assure que cette fonction atteint son minimum sur  $\left] 0, \frac{1}{a_d^*} \right]$ .

Finalement chaque  $\lambda_{\varepsilon_m}$  est tel que

$$\frac{1}{\lambda_{\varepsilon_m}} \in M_d(\varepsilon_m) \subset \left] 0, \frac{1}{a_d^*} \right]$$

$$\text{D'où } \lambda_{\varepsilon_m} \geq a_d^*$$

$$\text{et donc } \lambda_\infty \geq a_d^*$$

En combinant 1) et 2), nous obtenons (4.1.8)  $\square$

Le cas intéressant où l'on espérait que la limite de  $\lambda_d(\varepsilon)$  quand  $\varepsilon \rightarrow 0$  soit  $\{0\}$  se réalise dans le cadre du corollaire suivant.

### Corollaire 1

Les assertions suivantes sont équivalentes:

a)  $a_d^* = 0$

b)  $\Lambda_d(\varepsilon)$  est non vide pour  $\varepsilon$  dans un voisinage de 0 et  $\lim_{\varepsilon \rightarrow 0^+} \Lambda_d(\varepsilon) = \{0\}$

L'exemple de la fonction quadratique  $x^2$  illustre ce cas. En effet  $a_d^* = 0$  puisque la fonction  $f_d(x)$  n'est jamais affine sur  $\mathbb{R}_0^+$ . D'autre part, on vérifie aisément que la limite des  $\Lambda_d(\varepsilon)$  vaut 0.

$$\lim_{\varepsilon \rightarrow 0} \Lambda_d(\varepsilon) = \lim_{\varepsilon \rightarrow 0} \left\{ \sqrt{\frac{2\varepsilon}{\langle \Lambda_d, d \rangle}} \right\} = \{0\}.$$

Dans ce premier paragraphe, nous nous sommes attachés, grâce à l'introduction de  $a_d^*$ , à préciser la classification des différentes situations concernant  $\Lambda_d(\varepsilon)$  et à affirmer que la limite des intervalles  $\Lambda_d(\varepsilon)$  quand  $\varepsilon$  tend vers 0 est toujours le singleton  $\{a_d^*\}$  quand celui-ci est fini. Ceci nous permettra d'établir un développement en série de la fonction  $f'_\varepsilon(x_0; d)$  pour  $\varepsilon$  petit. Dans ce but, l'étude du comportement et la limite de  $f'_\varepsilon(x_0; d)$  s'avère nécessaire.



## IV.2 Etude de $f'_\varepsilon(x_0; d)$ en fonction de $\varepsilon$ .

---

Pour un point  $x_0$  et une direction  $d$  fixés, nous noterons la fonction  $\varepsilon \rightarrow f'_\varepsilon(x_0; d)$  par  $\theta_d(\cdot)$  que nous prolongerons sur  $\mathbb{R}$ . Le lien entre les fonctions  $\theta_d(\cdot)$  et  $\pi_{d,\varepsilon}(\cdot)$  est le point de départ à l'étude du sur-différentiel de la fonction  $\theta_d(\cdot)$ . Celle-ci nous permettra de dégager une propriété importante de l'ensemble  $\Pi_d(\varepsilon)$ . Ensuite, quelques théorèmes nous suggéreront un développement en série de  $\theta_d(\varepsilon)$ , dans le cas où  $0 < \alpha_d^* < +\infty$ , c'est quand  $f$  est affine sur un intervalle  $[x_0, x_0 + \lambda d]$  pour un certain  $\lambda \in \mathbb{R}_0^+$ .

### A. Définitions de $\theta_d(\cdot)$ et $\pi_{d,\varepsilon}(\cdot)$ sur $\mathbb{R}$ .

---

1)  $\theta_d : \mathbb{R} \rightarrow \mathbb{R}$

$$\varepsilon \mapsto \begin{cases} f'_\varepsilon(x_0; d) & \text{si } \varepsilon > 0 \\ -\infty & \text{sinon} \end{cases}$$

2)  $\pi_{d,\varepsilon} : \mathbb{R} \rightarrow \mathbb{R}$

$$\eta \mapsto \begin{cases} \eta \left[ f\left(x_0 + \frac{d}{\eta}\right) - f(x_0) + \varepsilon \right] & \text{pour } \eta > 0 \\ +\infty & \text{pour } \eta < 0 \\ f'_\infty(d) & \text{pour } \eta = 0 \end{cases}$$

### B) Propriétés de $\theta_d(\varepsilon)$

- on sait que  $\theta_d$  est une fonction croissante en  $\varepsilon$  et que  

$$\lim_{\varepsilon \rightarrow 0} \theta_d(\varepsilon) = \theta_d(0) = f'(x_0; d)$$
- Clarifions le lien entre  $\theta_d(\varepsilon)$  et  $\pi_{d,\varepsilon}(u)$ . Ceci mettra en évidence le caractère concave de  $\theta_d(\varepsilon)$ .

#### Proposition 4.2.1

$$(\pi_d)^*(y^*) = -\theta_d(-y^*) \text{ pour tout } y^* \in \mathbb{R}$$

Preuve.

Par définition de la fonction conjuguée,

$$\begin{aligned} (\pi_d)^*(y^*) &= \sup_{y \in \mathbb{R}} \{ y y^* - \pi_d(y) \} \\ &= \sup_{y > 0} \left\{ y y^* - y \left[ f\left(x_0 + \frac{d}{y}\right) - f(x_0) \right] \right\} \\ &\quad \text{car } \pi_d(y) = +\infty \text{ pour } y < 0 \end{aligned}$$

1) pour  $y^* > 0$

$$\pi_d^*(y^*) = +\infty \text{ car } \lim_{y \rightarrow \infty} \pi_d(y) = f'(x_0; d) < +\infty$$

$$-\theta_d(-y^*) = +\infty \text{ par définition de } \theta_d(\cdot) \text{ sur } \mathbb{R}_0^-$$

2) pour  $y^* \leq 0$

$$(\pi_d)^*(y^*) = - \inf_{y > 0} \left\{ y \left[ f\left(x_0 + \frac{d}{y}\right) - f(x_0) \right] + y(-y^*) \right\}$$

$$\begin{aligned}
 &= -f'_{(-y^*)}(x_0; d) \text{ par définition de } f'(x_0; d) \\
 &= -\theta_d(-y^*)
 \end{aligned}$$

$$\text{Donc } \forall y^* \quad (\pi_d)^*(y^*) = -\theta_d(-y^*)$$

□

Remarquons que, puisque  $\pi_{d,\varepsilon}(y) = \pi_d(y) + \varepsilon y$ , la proposition 4.2.1 associe l'expression de  $(\pi_{d,\varepsilon})^*(y^*)$ :

$$(\pi_{d,\varepsilon})^*(y^*) = -\theta_d(\varepsilon - y^*) \text{ pour tout } y^*$$

Comme toute fonction p-convexe est convexe, la fonction  $-\theta_d(\cdot)$  l'est et donc  $\theta_d(\cdot)$  est concave sur  $\mathbb{R}$ .

On peut dès lors déterminer son sur-différentiel, ce qui élucidera complètement le comportement qualitatif de  $f'_\varepsilon(x_0)$  pour  $\varepsilon > 0$ . Pour cela, explicitons tout d'abord le sous-différentiel de  $\pi_d$  sous forme de propositions.

#### Propositions 4.2.2

$$\text{i) } \partial \pi_d(0) = ]-\infty, \varepsilon_d^*]$$

$$\text{ii) } \forall \varepsilon > 0 \quad \partial \pi_{d,\varepsilon}(y) = -\frac{1}{\varepsilon} \langle \partial f(x_0 + \frac{d}{\varepsilon}), d \rangle + f(x_0 + \frac{d}{\varepsilon}) - f(x_0) + \varepsilon \quad \forall y > 0$$

$$\text{iii) } y_0 \neq 0, y_0 \in M_d(\varepsilon) \Leftrightarrow 0 \in \partial \pi_{d,\varepsilon}(y_0)$$

Énonçons et démontrons le résultat fondamental sur la structure du sur-différentiel de  $\theta_d(\varepsilon)$

### Théorème 4.2

$$i) \partial \theta_d(\varepsilon) = M_d(\varepsilon) \text{ pour tout } \varepsilon \geq 0$$

En particulier,  $\theta_d$  est strictement croissante sur  $[0, \varepsilon_d^*]$  et prend une valeur constante sur  $]\varepsilon_d^*, +\infty[$

$$ii) \pi_d(\eta) = \sup_{\varepsilon \geq 0} [J'_\varepsilon(x_0; d) - \varepsilon \eta] \text{ pour tout } \eta \geq 0$$

### Preuve de i)

Le sur-différentiel de la fonction  $\theta_d$  en  $\varepsilon$  est l'opposé du sous-différentiel de  $(-\theta_d)$  en  $\varepsilon$ , c.-à.-d

$$\partial \theta_d(\varepsilon) = -\partial(-\theta_d)(\varepsilon) \quad \forall \varepsilon \geq 0.$$

La proposition 4.2.1 nous renseigne que  $(\pi_d)^*(\eta^*) = -\theta_d(-\eta^*)$  et nous permet de tirer les équivalences suivantes pour tout  $\varepsilon \geq 0$

$$\begin{aligned} \eta^* \in \partial \theta_d(\varepsilon) &\Leftrightarrow -\eta^* \in \partial(-\theta_d)(\varepsilon) \\ &\Leftrightarrow \eta^* \in \partial(\pi_d^*)(-\varepsilon) \\ &\Leftrightarrow -\varepsilon \in \partial \pi_d(\eta^*) \quad [\text{II.2 remarque 1}] \end{aligned}$$

• si  $\eta^* = 0$

$$0 \in \partial \theta_d(\varepsilon) \Leftrightarrow -\varepsilon \in \partial \pi_d(0) = ]-\infty, \varepsilon_d^*] \text{ par la proposition 4.22.}$$



$$\Leftrightarrow \varepsilon \geq \varepsilon_d^*$$

$$\Leftrightarrow 0 \in M_d(\varepsilon) \quad [\text{IV.1 (A2)}]$$

• si  $u^* > 0$

$$u^* \in \partial \theta_d(\varepsilon) \Leftrightarrow -\varepsilon \in \partial \pi_d(u^*)$$

$$\Leftrightarrow 0 \in \partial \pi_d(u^*) + \varepsilon$$

$$\Leftrightarrow 0 \in -\frac{1}{u^*} \langle \partial f(x_0 + \frac{d}{u^*}), d \rangle + f(x_0 + \frac{d}{u^*}) - f(x_0) + \varepsilon$$

$$\Leftrightarrow 0 \in \partial \pi_{d,\varepsilon}(u^*)$$

$$\Leftrightarrow u^* \text{ minimise } \pi_{d,\varepsilon}(u) \text{ sur } \mathbb{R}^+$$

$$\Leftrightarrow u^* \in M_d(\varepsilon)$$

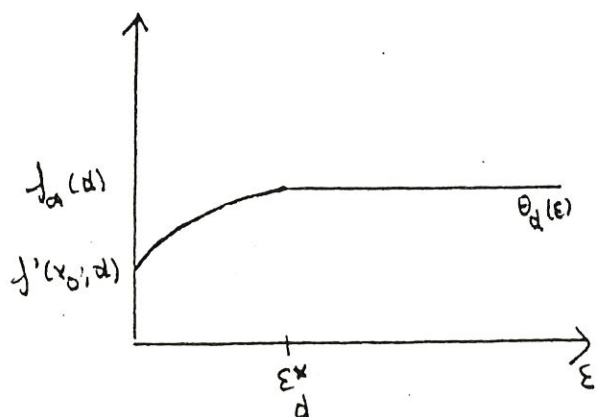
Dès lors, on a  $\partial \theta_d(\varepsilon) = M_d(\varepsilon) \quad \forall \varepsilon \geq 0$ .

Pour prouver la stricte croissance de  $\theta_d(\cdot)$  sur  $[0, \varepsilon_d^*]$ , il suffit de constater que  $0 \notin \partial \theta_d(\varepsilon)$  pour  $0 < \varepsilon < \varepsilon_d^*$ , puisque l'on sait que  $\theta_d(\cdot)$  est croissante sur  $\mathbb{R}^+$ . De plus, sur  $[\varepsilon_d^*, +\infty[$ ,  $\theta_d(\cdot)$  est constante et vaut  $f_\infty(d)$ .

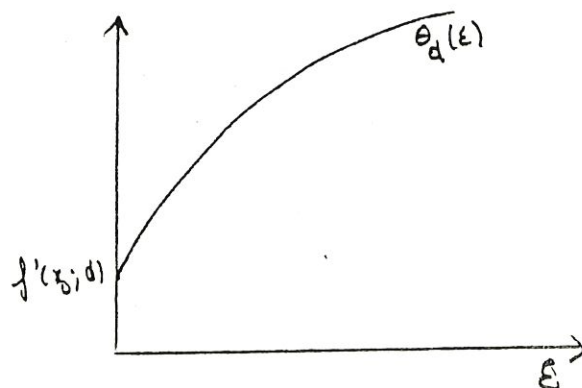
□

Représentons les deux situations possibles du graphe de  $\theta_d(\varepsilon)$  suivant la valeur de  $f_\infty(d)$

$$f_\infty(d) < +\infty$$



$$f_\infty(d) = +\infty \Leftrightarrow \varepsilon_d^* = +\infty$$



Preuve de ii)

Calculons pour tout  $\mu \geq 0$  le  $\sup_{\varepsilon > 0} [f'_\varepsilon(x_0; d) - \varepsilon \mu]$  et montrons que'il vaut  $\pi_d(\mu)$ .

$$\begin{aligned} \sup_{\varepsilon > 0} [f'_\varepsilon(x_0; d) - \varepsilon \mu] &= \sup_{\varepsilon} [\theta_d(\varepsilon) - \varepsilon \mu] \text{ car } \theta_d(\varepsilon) = f'_\varepsilon(x_0; d) \\ &= \sup_{\varepsilon} [-\pi_d^*(\varepsilon) - \varepsilon \mu] \text{ par la proposition 4.2.1} \\ &= \sup_{\varepsilon} [-\pi_d^*(\varepsilon) + \varepsilon \mu] \\ &= \pi_d^{**}(\mu) \text{ par définition de la bidualité} \\ &= \pi_d(\mu) \text{ car } \pi_d(\cdot) \text{ est convexe.} \end{aligned}$$

□

Remarque

La connaissance de  $\lambda_d(\varepsilon)$  et celle de  $\theta_d$  (ou, de manière équivalente, celle de  $\mu_d(\varepsilon)$  et  $\pi_d$ ) permettent de calculer  $f'_\varepsilon(x_0; d)$  puisque

$$f'_\varepsilon(x_0; d) = \frac{f(x_0 + \lambda d) - f(x_0) + \varepsilon}{\lambda} = \theta_d(\lambda) + \frac{\varepsilon}{\lambda} \text{ pour } \lambda \in \lambda_d(\varepsilon)$$

Réciproquement, si  $f'_\varepsilon(x_0; d)$  est connu pour tout  $\varepsilon > 0$ ,

la relation ii) fournit l'expression de  $\pi_d$  sur  $\mathbb{R}_+$

puisque : 
$$\pi_d(\mu) = \sup_{\varepsilon > 0} [f'_\varepsilon(x_0; d) - \varepsilon \mu] \text{ pour tout } \mu \geq 0.$$

Illustrons ce dernier théorème sur deux exemples.

exemple 10

Soit  $f: \mathbb{R} \longrightarrow \mathbb{R}$

$$x \longmapsto \begin{cases} 2x^{3/2} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$$

Fixons  $x_0 = 0$ ,  $d = 1$ .

Pour vérifier que  $\partial \theta_d(\varepsilon) = M_d(\varepsilon)$ , calculons d'abord  $q_{d,\varepsilon}(\lambda)$ ,  $f'_\varepsilon(x_0; d)$  et  $\Lambda_d(\varepsilon)$ .

$$q_{d,\varepsilon}(\lambda) = 2\sqrt{\lambda} + \frac{\varepsilon}{\lambda}$$

Cette fonction atteint son minimum  $3\varepsilon^{1/3}$  en  $\lambda = \varepsilon^{2/3}$ .

$$\text{Donc lors } f'_\varepsilon(x_0; d) = \theta_d(\varepsilon) = 3\varepsilon^{1/3}$$

$$\Lambda_d(\varepsilon) = \{\varepsilon^{2/3}\}$$

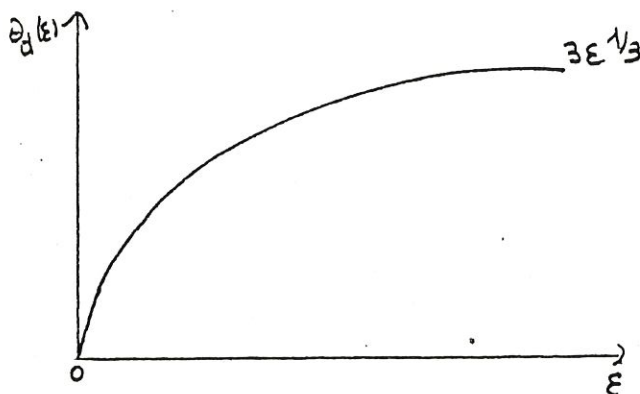
$$M_d(\varepsilon) = \{\varepsilon^{-2/3}\}$$

Calculons  $\partial \theta_d(\varepsilon)$

$$\partial \theta_d(\varepsilon) = \left\{ \frac{d}{d\varepsilon}(3\varepsilon^{1/3}) \right\} = \{\varepsilon^{-2/3}\}.$$

Nous pouvons donc vérifier que  $\partial \theta_d(\varepsilon) = M_d(\varepsilon)$

Remarquons de plus que  $\theta_d(\varepsilon)$  est strictement croissante sur  $\mathbb{R}^+$ . Ceci est naturel puisque  $f_\infty(d) = +\infty$



Grâce à la connaissance de la fonction  $\theta_d(\varepsilon)$ , nous pouvons retrouver celle de  $r_d(\eta)$

$$\text{En effet } r_d(\eta) = \sup_{\varepsilon > 0} \{3\varepsilon^{1/3} - \varepsilon\eta\} = \frac{2}{\sqrt{\eta}} = q_d\left(\frac{1}{\eta}\right)$$

A l'opposé de l'exemple 10, considérons maintenant une fonction où  $f_\infty(d)$  est fini. La fonction polyédrale de l'exemple 9, rappelée ci-dessous, satisfait cette condition

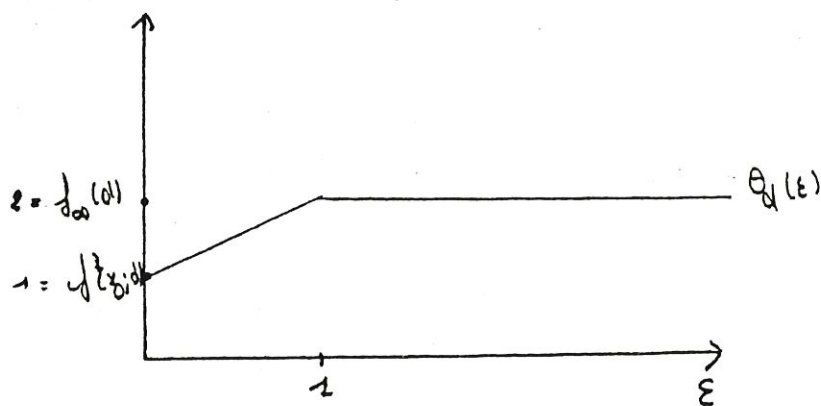
$$f: \mathbb{R} \longrightarrow (-\infty, +\infty]$$

$$f(x) = \begin{cases} x & \text{si } 0 \leq x \leq 1 \\ 2x-1 & \text{si } x > 1 \\ +\infty & \text{si } x < 0 \end{cases}$$

D'une part nous avons déjà calculé que  $\varepsilon_d^* = 1$

D'autre part, on sait que  $\Theta_d(\varepsilon) = 1 + \varepsilon \text{ sur } [0, 1]$   
 $\quad \quad \quad = 2 = f_\infty(d) \text{ sur } [1, +\infty[$

$\Theta_d(\cdot)$  est donc strictement croissante sur  $[0, \varepsilon_d^*]$  et est constante sur  $[\varepsilon_d^*, +\infty[$



Suite au dernier théorème, nous pouvons déduire une conséquence importante en particulierisant  $M_d(\varepsilon)$  pour presque tous les  $\varepsilon > 0$ .



### Proposition 4.2.3

- i)  $M_d(\varepsilon)$  est un singleton presque partout  
 ii)  $\Lambda_d(\varepsilon)$  est un singleton presque partout  
 sur  $]0, \varepsilon_d^* [$ .

#### Preuve

i) Comme  $\Theta_d(\varepsilon)$  est concave, elle est différentiable p.p. [chap. (4.7)]. Son sur-différentiel est donc réduit à un singleton p.p. Dès lors  $M_d(\varepsilon) (= \partial \Theta_d(\varepsilon))$  est un singleton p.p.

ii) Puisque  $\Lambda_d(\varepsilon) = \emptyset \quad \forall \varepsilon > \varepsilon_d^*$  et  $d_d(\varepsilon) = \frac{1}{M_d(\varepsilon)} \quad \forall \varepsilon \in ]0, \varepsilon_d^* [$ , la fonction  $q_{d,\varepsilon}$  atteint sa valeur minimale en un seul point  $\lambda_d(\varepsilon)$ , sauf en un nombre fini de  $\lambda$  et donc  $\Lambda_d(\varepsilon)$  est un singleton p.p. sur  $]0, \varepsilon_d^* [$ .

□

À côté du comportement qualitatif de  $f'_\varepsilon(x_0; d)$  en fonction de  $\varepsilon$ , il est intéressant d'étudier le comportement limite du quotient  $[f'_\varepsilon(x_0; d) - f'(x_0; d)] / \varepsilon$  quand  $\varepsilon \rightarrow +\infty$  ou  $\varepsilon \rightarrow 0$ .

c) Comportement de  $f'_\varepsilon(x_0; d)$  quand  $\varepsilon \rightarrow +\infty$ .

---

On sait que,  $\forall \varepsilon \geq \varepsilon_d^* \quad f'_\varepsilon(x_0; d) = f'_\infty(d)$ .

Quand  $f_\infty(d) = +\infty$ , on a donc que  $\lim_{\varepsilon \rightarrow +\infty} f'_\varepsilon(x_0; d) = +\infty$

On pourrait être tenté de décrire cette convergence comme celle d'une application linéaire en  $\varepsilon$ . En fait, il n'en est rien.

#### Proposition 4.2.4

$$\begin{array}{l} 1) \lim_{\varepsilon \rightarrow +\infty} \frac{f'_\varepsilon(x_0; d)}{\varepsilon} = 0 \\ 2) \lim_{\varepsilon \rightarrow +\infty} M_d(\varepsilon) = \{0\} \end{array}$$

Ce résultat, démontré dans ([3], p. 14), n'a de portée que si  $f_\infty(d) = +\infty$ . Dans le cas contraire, on sait en effet que  $f'_\varepsilon(x_0; d) = f'_\infty(d)$  et  $M_d(\varepsilon) = \{0\}$  pour tout  $\varepsilon > \varepsilon_d^*$ . Examinons donc sur 1° exemple le cas intéressant.

En effet,  $f_\infty(d) = +\infty$

$$\text{et } \lim_{\varepsilon \rightarrow +\infty} \frac{f'_\varepsilon(x_0; d)}{\varepsilon} = \lim_{\varepsilon \rightarrow +\infty} 3 \cdot \frac{\varepsilon^{1/3}}{\varepsilon} = 0.$$

$$\lim_{\varepsilon \rightarrow +\infty} M_d(\varepsilon) = \lim_{\varepsilon \rightarrow +\infty} \{\varepsilon^{-2/3}\} = \{0\}$$

Le comportement limite de  $[f'_\varepsilon(x_0; d) - f'_\infty(x_0; d)]/\varepsilon$  quand  $\varepsilon \rightarrow 0^+$  est plus important à appréhender.

D) Comportement de  $f'_\varepsilon(x_0; d)$  quand  $\varepsilon \rightarrow 0^+$

---

Pour élucider cette variation, énonçons et démontrons le théorème suivant.

### Théorème 4.3

$$\Theta'_d(0) = \lim_{\varepsilon \rightarrow 0} \frac{f'_\varepsilon(x_0; d) - f'(x_0; d)}{\varepsilon} = \frac{1}{a_d^*} \text{ dans } \overline{\mathbb{R}}_+$$

#### Preuve

$$(\Theta_d)'_+(0) = \lim_{\varepsilon \rightarrow 0^+} \frac{\Theta_d(\varepsilon) - \Theta_d(0)}{\varepsilon} = \Theta'_d(0; 1)$$

où  $\Theta'_d(0; 1)$ , dérivée directionnelle de  $\Theta_d$  dans la direction 1, est encore la fonction d'appui du sur-différentiel de la fonction concave  $\Theta_d$  au point 0 dans la direction 1

$$\text{Donc } (\Theta_d)'_+(0) = \min_{\eta_0 \in \partial \Theta_d(0)} \langle \eta_0, 1 \rangle$$

$$\eta_0 \in \partial \Theta_d(0)$$

$$= \min \partial \Theta_d(0)$$

$$= \min M_d(0) \text{ car } \partial \Theta_d(0) = M_d(0)$$

$$= \frac{1}{a_d^*} \text{ car } M_d(0) = \left[ \frac{1}{a_d^*}, +\infty \right]$$

□

Le théorème 4.3 est un outil opportun à l'élaboration d'un développement en série de  $f'_\varepsilon(x_0; d)$ . Pour cela considérons les deux cas suivants.

$$1) \underline{0 < a_d^* < +\infty}$$

Il découle immédiatement du théorème 4.3 que

$$f'_\varepsilon(x_0; d) = f'(x_0; d) + \frac{\varepsilon}{a_d^*} + o(\varepsilon) \quad (4.2.1)$$

En outre, le théorème 4.1 nous permet de déduire que

$$f'_\varepsilon(x_0; d) = f'(x_0; d) + \frac{\varepsilon}{\lambda_d(\varepsilon)} + o(\varepsilon) \quad \forall \lambda_d(\varepsilon) \in \Lambda_d(\varepsilon)$$

La situation typique qui satisfait cette hypothèse est le cas où  $f$  est polyédrale mais pas affine sur  $[x_0, x_0 + R^+d]$  (exemples). En conséquence, la fonction  $\theta_d(\varepsilon)$  est également polyédrale sur  $R^+$ . Le terme  $o(\varepsilon)$  disparaît donc dans l'égalité (4.21).

En résumé, si  $0 < a_d^* < +\infty$  et si  $f$  est polyédrale sur  $R^+$  alors  $f'_\varepsilon(x_0; d) = f'(x_0; d) + \frac{\varepsilon}{a_d^*}$  quand  $\varepsilon$  est petit.

Reprenons l'exemple où  $f: R \rightarrow \bar{R}$

$$x \mapsto \begin{cases} 2x-1 & \text{si } x \geq 1 \\ x & \text{si } 0 \leq x \leq 1 \\ +\infty & \text{ailleurs} \end{cases}$$

Nous avons  $\varepsilon_d^* = 1$

$a_d^* = 1$

$$\left. \begin{aligned} f'_\varepsilon(x_0; d) &= 1 + \varepsilon & \text{si } \varepsilon \leq \varepsilon_d^* \\ &= 2 = f_\infty(d) & \text{si } \varepsilon \geq \varepsilon_d^* \end{aligned} \right\}$$

$f'(x_0; d) = 1$

On vérifie que  $f'_\varepsilon(x_0; d) = 1 + \varepsilon = f'(x_0; d) + \frac{\varepsilon}{a_d^*}$

2)  $a_d^* = 0$

Dans ce cas, un tel développement ne subsiste pas car

les quantités  $\frac{f'_\varepsilon(x_0; d) - f'(x_0; d)}{\varepsilon}$  et  $\frac{1}{\lambda_d(\varepsilon)}$  tendent toutes deux



vers l'infini quand  $\varepsilon$  tend vers 0 et leur différence ne s'annule pas nécessairement à la limite. Toutefois, on a le résultat plus faible suivant.

Proposition 4.2.5.

Supposons que  $a_d^* < +\infty$

Alors  $\lim_{\varepsilon \rightarrow 0^+} \left[ f'_\varepsilon(x_0; d) - f'(x_0; d) - \frac{\varepsilon}{\lambda_d(\varepsilon)} \right] = 0$

quelque soit  $\lambda_d(\varepsilon)$  dans  $\Lambda_d(\varepsilon)$

Preuve

Soit  $\lambda_d(\varepsilon) \in \Lambda_d(\varepsilon)$ .

$\frac{1}{\lambda_d(\varepsilon)}$  est un sub-gradient de  $\Theta_d$  en  $\varepsilon > 0$  car  $\partial \Theta_d(\varepsilon) = N_d(\varepsilon) = \frac{1}{\lambda_d(\varepsilon)}$

Dès lors, nous avons :

$$f'_\varepsilon(x_0; d) - f'(x_0; d) \geq \frac{\varepsilon}{\lambda_d(\varepsilon)} \quad \forall \varepsilon \in ]0, \varepsilon^*]$$

Comme le premier membre de l'inégalité tend vers 0 quand  $\varepsilon$  tend vers 0,  $\frac{\varepsilon}{\lambda_d(\varepsilon)}$  aussi et il découle :

$$\lim_{\varepsilon \rightarrow 0^+} \left[ f'_\varepsilon(x_0; d) - f'(x_0; d) - \frac{\varepsilon}{\lambda_d(\varepsilon)} \right] = 0$$

□

L'exemple de la fonction  $x \mapsto |x|^d$  avec  $d \in ]1, 2[$ ,  $x_0 = 0$  et  $d = 1$  réalise la condition  $a_d^* = 0$ . Il montre bien que, malgré la valeur infinie des limites des expressions

$\frac{f'_\varepsilon(x_0; d) - f'(x_0; d)}{\varepsilon}$  et  $\frac{1}{\lambda_d(\varepsilon)}$ , leur différence ne tend pas nécessairement vers 0. On constatera toutefois le résultat affaibli de la proposition 4.2.5.

### Exemple 4.1

Soit  $f: \mathbb{R} \rightarrow \mathbb{R}$

$$x \mapsto |x|^d \quad \text{où } d \in ]1, 2[$$

Choisissons  $x_0 = 0$ ,  $d = 1$

Calculons d'abord  $a_d^*$ ,  $f'_\varepsilon(x_0; d)$  et  $\lambda_d(\varepsilon)$

•  $q_d(\lambda) = \lambda^{d-1}$  et dès lors  $f'(x_0; d) = 0$   
 comme  $q_d(\lambda) > f'(x_0; d) \quad \forall \lambda > 0$ ,  $a_d^* = 0$

•  $q_{d,\varepsilon}(\lambda) = \lambda^{d-1} + \frac{\varepsilon}{\lambda}$   
 $f'_\varepsilon(x_0; d) = \inf_{\lambda > 0} q_{d,\varepsilon}(\lambda) = \alpha \left( \frac{\varepsilon}{d-1} \right)^{\frac{d-1}{d}} \quad \forall \varepsilon > 0$

$$\lambda_d(\varepsilon) = \left\{ \left( \frac{\varepsilon}{d-1} \right)^{1/d} \right\}$$

On remarque que  $\lim_{\varepsilon \rightarrow 0} \frac{f'_\varepsilon(x_0; d) - f'(x_0; d)}{\varepsilon} = +\infty$

et  $\lim_{\varepsilon \rightarrow 0} \frac{1}{\lambda_d(\varepsilon)} = +\infty$ .

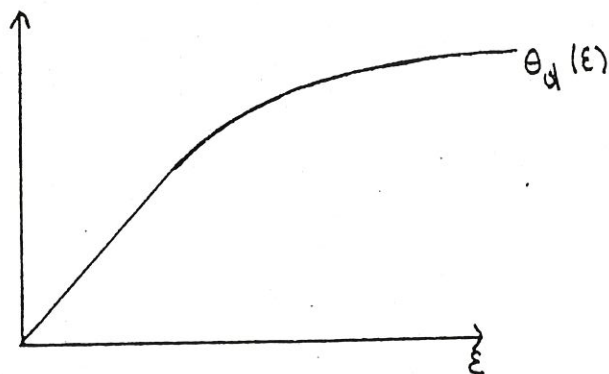
La différence  $\frac{f'_\varepsilon(x_0; d) - f'(x_0; d)}{\varepsilon} - \frac{1}{\lambda_d(\varepsilon)}$  se comporte comme  $\varepsilon^{-1/d}$  et donc tend vers l'infini quand  $\varepsilon$  tend vers 0.

Cependant on a le résultat affaibli:  $f'_\varepsilon(x_0; d) - f'(x_0; d) - \frac{\varepsilon}{\lambda_d(\varepsilon)}$  se comporte en  $\varepsilon^{\frac{d-1}{d}}$  et donc tend vers 0 avec  $\varepsilon$  (pour  $d \in ]1, 2[$ )

Schématisons l'apport du théorème 4.3 par des graphes de  $\theta_d(\cdot)$  suivant les valeurs de  $a_d^*$ .

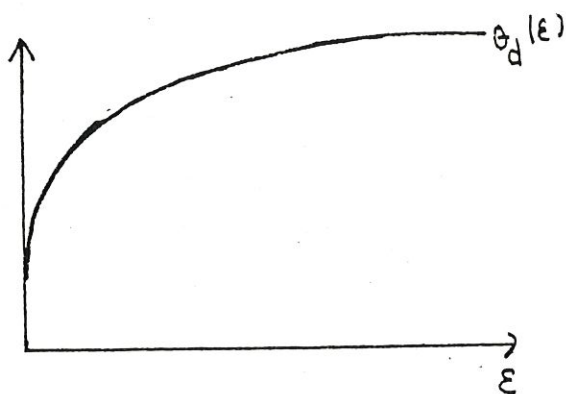
1) Si  $0 < a_d^* < +\infty$

Le théorème nous renseigne que la pente de la tangente à  $\theta_d(\cdot)$  en 0 est finie:  $(\theta_d')_+(0) = \frac{1}{a_d^*}$



2) Si  $a_d^* = 0$

Dans ce cas  $(\theta_d')_+(0) = \infty$  et la tangente en 0 est verticale



Nous avons montré dans ce paragraphe que  $\frac{\epsilon}{\lambda_d(\epsilon)}$  tend vers 0 avec  $\epsilon$ . Sans information supplémentaire sur le comportement de  $f_d$  au voisinage de  $0^+$ , on ne peut

assurer que  $\frac{\varepsilon}{(\lambda_d(\varepsilon))^\beta}$  converge vers 0 pour  $\beta > 1$ . Ceci montre que les résultats du théorème 4.3 et de la proposition 4.2.5 sont "optimaux". Enfin, en règle générale, on ne peut espérer trouver  $\sigma > 0$  tel que

$$f''_\varepsilon(x_0; d) - f'(x_0; d) = O(\varepsilon^\sigma).$$

L'exemple précédent nous en convainc déjà:

D'une part,

$$\lim_{\varepsilon \rightarrow 0} \frac{\varepsilon}{(\lambda_d(\varepsilon))^\beta} = \lim_{\varepsilon \rightarrow 0} \varepsilon^{1-\beta/d} (d-1)^{\beta/d} = +\infty \quad \forall \beta > 2$$

D'autre part, puisque  $a_d^* = 0$ , on ne peut trouver  $\sigma$  tel que  $f''_\varepsilon(x_0; d) - f'(x_0; d) = d \left( \frac{\varepsilon}{d-1} \right)^{\frac{d-1}{d}} = O(\varepsilon^\sigma)$ . En effet, ceci n'est pas possible pour  $d > 1$ .

Remarquons que, pour  $a_d^* > 0$ ,  $\sigma = 1$  convient puisque

$$\lim_{\varepsilon} \frac{f''_\varepsilon(x_0; d) - f'(x_0; d)}{\varepsilon} = \frac{1}{a_d^*} < +\infty$$

Grâce à cette étude du comportement en  $\varepsilon$  de la dérivée directionnelle approchée  $f''_\varepsilon(x_0; d)$ , pour un point  $x_0$  et une direction  $d$  fixés, nous sommes prêts à examiner une tentative de réponse à la question posée dès le début de ce mémoire: dans quelle mesure  $f''_\varepsilon(x_0; d; \delta)$  est-elle une approximation de la dérivée seconde de  $f$  en  $x_0$ ?



### IV.3. Comportement de $f''_{\varepsilon}(x_0; d, \delta)$ comme fonction de $\varepsilon$ .

---

Dans ce paragraphe, nous centrerons notre attention sur le comportement limite de  $f''_{\varepsilon}(x_0; d, \delta)$  quand  $\varepsilon$  tend vers 0. Nous débuterons cette étude grâce à la distinction déjà abordée au premier paragraphe suivant que  $a_d^*$  vaut l'infini, zéro ou un nombre fini non nul. Ce dernier cas nous fournira un développement en série du premier ordre en  $\varepsilon$  pour  $f''_{\varepsilon}(x_0; d, \delta)$ . La deuxième partie sera consacrée à  $f''_{\varepsilon}(x_0; d, \delta)$  comme approximation de la dérivée de Dini en  $x_0$ . Signalons dès à présent que les résultats les plus probants seront obtenus dans le cas diagonal.

#### A. Etude de $f''_{\varepsilon}(x_0; d, \delta)$ axée sur la valeur de $a_d^*$

---

Rappelons que  $f''_{\varepsilon}(x_0; d, \delta)$ , comme dérivée directionnelle de  $f'_{\varepsilon}(x_0; d)$  en  $x_0$  dans la direction  $\delta$ , s'écrit comme suit:

$$f''_{\varepsilon}(x_0; d, \delta) = \lim_{\delta \rightarrow 0} \frac{f'_{\varepsilon}(x_0 + \delta; d) - f'_{\varepsilon}(x_0; d)}{\delta} \quad \text{et par le théorème 3.1}$$

$$= \min_{\eta \in M_{d, \varepsilon}(x_0)} \left\{ \eta \left[ \left\langle \Psi_{\frac{\partial f(x_0)}{\partial \varepsilon}}^*(\delta) - f'(x_0; \delta) \right\rangle \right] \right\} \quad (4.3.1)$$

où  $\psi_{\frac{\gamma}{\varepsilon} f(x_0)_D}^*(\delta)$  est la fonction d'appui de  $\frac{\gamma}{\varepsilon} f(x_0)_D$

avec  $\partial_{\frac{\gamma}{\varepsilon} f(x_0)_D} = \{x^* \in \partial_{\frac{\gamma}{\varepsilon} f(x_0)} \text{ t.q. } \langle x^*, d \rangle = f'_{\frac{\gamma}{\varepsilon}}(x_0; d)\}$

Quand  $\Lambda_d(\varepsilon)$  est non vide et borné,  $f''_{\frac{\gamma}{\varepsilon}}(x_0; d; \delta)$  peut s'exprimer comme :

$$f''_{\frac{\gamma}{\varepsilon}}(x_0; d; \delta) = \min_{\lambda \in \Lambda_d(\varepsilon)} \left\{ \frac{\psi_{\frac{\gamma}{\varepsilon} f(x_0)_D}^*(\delta) - f'_{\frac{\gamma}{\varepsilon}}(x_0; \delta)}{\lambda} \right\}$$

Suivant que  $[\psi_{\frac{\gamma}{\varepsilon} f(x_0)_D}^*(\delta) - f'_{\frac{\gamma}{\varepsilon}}(x_0; \delta)]$  est positif ou négatif, le minimum sur  $\Lambda_d(\varepsilon)$  est atteint pour  $\bar{\lambda}_{\varepsilon}$  ou  $\underline{\lambda}_{\varepsilon}$  qui sont respectivement le maximum et le minimum de l'ensemble  $\Lambda_d(\varepsilon)$ .

Dans le cas diagonal, nous avons :

$$f''_{\frac{\gamma}{\varepsilon}}(x_0; d; d) = \frac{f'_{\frac{\gamma}{\varepsilon}}(x_0; d) - f'(x_0; d)}{\bar{\lambda}_{\varepsilon}}$$

Dans le cas où  $\Lambda_d(\varepsilon)$  est non vide et borné, la fonction  $\varepsilon \rightarrow f''_{\frac{\gamma}{\varepsilon}}(x_0; d; d)$  est le quotient de deux fonctions décroissantes en  $\varepsilon$  et, par la proposition 4.11, nous avons :

$$f''_{\frac{\gamma}{\varepsilon}}(x_0; d; d) \geq \frac{\varepsilon}{(\bar{\lambda}_d(\varepsilon))^2} \quad (4.3.2)$$

$f''_{\frac{\gamma}{\varepsilon}}(x_0; d; d)$  peut être non bornée quand  $\varepsilon \rightarrow 0^+$  car on n'est pas assuré de la convergence du second membre de

l'inégalité (4.3.2) vers un nombre fini. De plus, il m'est pas espéré que  $f''_{\varepsilon}(x_0; d, \delta)$  soit croissante ou décroissante en  $\varepsilon$ . Dans le but d'étudier le comportement de  $f''_{\varepsilon}(x_0; d, \delta)$  quand  $\varepsilon \rightarrow 0$ , nous distinguons trois cas.

1)  $a_d^* = +\infty$

Dans un tel cas, on voit que  $f''_{\varepsilon}(x_0; d, \delta) = 0$  pour tout  $\varepsilon$  et pour toute direction  $\delta$ . En effet,  $\Lambda_d(\varepsilon) = \emptyset$  et  $M_d(\varepsilon) = \{0\}$  pour tout  $\varepsilon$ . Dès lors l'expression (4.3-1) s'annule et nous pouvons écrire.

$$\lim_{\varepsilon \rightarrow 0} f''_{\varepsilon}(x_0; d, \delta) = 0 \quad \forall \delta$$

2)  $0 < a_d^* < +\infty$

Dans ce cas, nous savons que  $\Lambda_d(\varepsilon)$  est un intervalle compact pour tout  $\varepsilon < \varepsilon_d^*$ . Comme nous nous intéressons au comportement de  $f''_{\varepsilon}$  pour  $\varepsilon$  petit, nous pouvons considérer les  $\varepsilon < \varepsilon_d^*$  et écrire

$$f''_{\varepsilon}(x_0; d, \delta) = \frac{4 \frac{1}{2} f(x_0; \delta) - f'(x_0; \delta)}{\lambda_{\varepsilon}}$$

où  $\lambda_{\varepsilon}$  est soit  $\lambda_{\varepsilon}$ , soit  $\bar{\lambda}_{\varepsilon}$  suivant que le numérateur est négatif ou positif. Le théorème 4.1 nous confirme la convergence de  $\lambda_{\varepsilon}$  vers  $a_d^*$ . L'existence de la limite de  $f''_{\varepsilon}(x_0; d, \delta)$  est donc équivalente à celle du numérateur.

Examinons dès lors  $\lim_{\varepsilon \rightarrow 0} \psi_{\partial_\varepsilon f(x_0)_d}^*(\delta)$  pour tout  $\delta$ .

Par la semi-continuité inférieure de la fonction  $\varepsilon \mapsto \psi_{\partial_\varepsilon f(x_0)_d}^*$ , on a

$$\limsup_{\varepsilon \rightarrow 0^+} \psi_{\partial_\varepsilon f(x_0)_d}^*(\delta) \leq \psi_{\partial f(x_0)_d}^*(\delta) \quad (4.3.3).$$

D'autre part

$$\liminf_{\varepsilon \rightarrow 0^+} \psi_{\partial_\varepsilon f(x_0)_d}^*(\delta) \geq \psi_{\partial f(x_0)_d}^*(\delta) \geq -\psi_{\partial f(x_0)_d}^*(-\delta) \quad (4.3.4)$$

Sous l'hypothèse :  $\langle x^*, \delta \rangle = \alpha$  constante  $\forall x^* \in \partial f(x_0)_d$  (\*)

$\lim_{\varepsilon \rightarrow 0} f''_\varepsilon(x_0; d, \delta)$  existe et est négative ou nulle.

Preuve

Soit une direction  $\delta$  telle que  $\langle x^*, \delta \rangle = \alpha \quad \forall x^* \in \partial f(x_0)_d$

$$\text{Alors } \psi_{\partial f(x_0)_d}^* = \max_{x^* \in \partial f(x_0)_d} \langle x^*, \delta \rangle = \alpha$$

$$\text{et } -\psi_{\partial f(x_0)_d}^*(-\delta) = \alpha$$

Donc, en remplaçant dans (4.3.3) et (4.3.4), on obtient :

$$\limsup_{\varepsilon \rightarrow 0^+} \psi_{\partial_\varepsilon f(x_0)_d}^*(\delta) \leq \alpha \leq \liminf_{\varepsilon \rightarrow 0^+} \psi_{\partial_\varepsilon f(x_0)_d}^*(\delta).$$

Comme la limite inférieure est toujours plus petite ou égale, à la limite supérieure, on déduit que

$$\lim_{\varepsilon \rightarrow 0} \psi_{\partial_\varepsilon f(x_0)_d}^* = \alpha \quad \text{et dès lors}$$

$$\lim_{\varepsilon \rightarrow 0} f''_\varepsilon(x_0; d, \delta) = \frac{\alpha - f'(x_0; \delta)}{\alpha^2 d} \quad (4.3.5)$$



Puisque  $f'(x_0; d) > d$  et  $a_d^*(x_0) > 0$ , cette limite est bien négative ou nulle

□

En particulier, le cas diagonal renforce l'hypothèse (\*) et la constante  $d$  vaut  $f'(x_0; d)$ . De (4.2.5), on peut alors déduire que

$$\lim_{\varepsilon \rightarrow 0} f''_{\varepsilon}(x_0; d, d) = 0$$

En résumé

$$\begin{aligned} &\text{Si } d \text{ est une direction telle que } 0 < a_d^* \leq +\infty. \\ &\text{Alors } \lim_{\varepsilon \rightarrow 0} f''_{\varepsilon}(x_0; d, d) = 0 \quad (4.3.6) \end{aligned}$$

On pourrait donc s'attendre à un développement en  $\varepsilon$  pour  $f''_{\varepsilon}(x_0; d, d)$ . En fait, il n'existe que dans le cas où celui de  $f'_{\varepsilon}(x_0; d)$  existe, c.-à-d. quand  $0 < a_d^* < +\infty$ .

$$\text{Si } 0 < a_d^* < +\infty$$

$$\text{Alors } f''_{\varepsilon}(x_0; d, d) = \frac{\varepsilon}{(a_d^*)^2} + o(\varepsilon) \quad (4.3.7)$$

Preuve

Pour  $0 < a_d^* < +\infty$ , nous avons par (4.2.1) que

$$f'_{\varepsilon}(x_0; d) = f'(x_0; d) + \frac{\varepsilon}{a_d^*} + o(\varepsilon)$$

Puisque  $f''_{\varepsilon}(x_0; d, d) = \frac{f'_{\varepsilon}(x_0; d) - f'(x_0; d)}{\lambda_{\varepsilon}}$  et que  $\{\lambda_{\varepsilon}\}_{\varepsilon}$  converge vers  $a_d^*$  quand  $\varepsilon \rightarrow 0^+$ , on obtient le résultat (4.3.7).

□

### Remarque

Le développement (4.3.7) est équivalent à

$$f''_{\varepsilon}(x_0; d, d) = \frac{\varepsilon}{(\lambda_{\varepsilon})^2} + o(\varepsilon) \quad \text{où } \lambda_{\varepsilon} \text{ est choisi arbitrairement dans } \Lambda_d(\varepsilon).$$

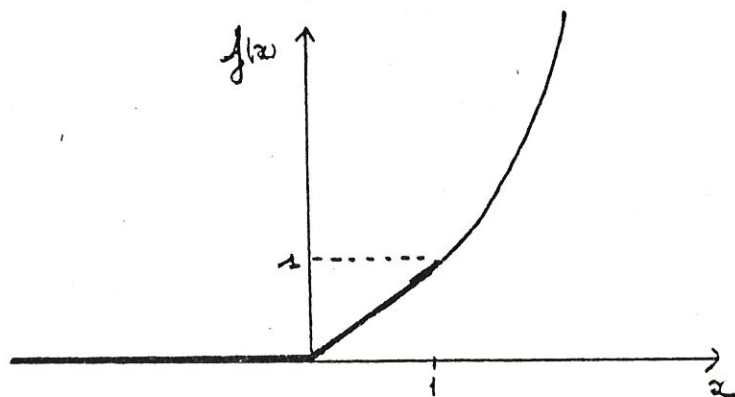
Voici maintenant un exemple où  $d$  et  $x_0$  sont tels que  $0 < a_d^* < +\infty$  et vérifient  $\lim_{\varepsilon \rightarrow 0} f''_{\varepsilon}(x_0; d, d) = 0$ .

### Exemple 12

Soit  $f: \mathbb{R} \rightarrow \mathbb{R}$

$$x \mapsto \begin{cases} 0 & \text{si } x \leq 0 \\ x & \text{si } 0 \leq x \leq 1 \\ x^2 & \text{si } x > 1 \end{cases}$$

Choisissons  $x_0 = 0$  et  $d = 1$



Calculons d'abord  $a_d^*$

$$q_d(\lambda) = \begin{cases} 1 & \text{si } \lambda \leq 1 \\ \lambda & \text{si } \lambda > 1 \end{cases}$$

Des lors  $f'(x_0; d) = 1$

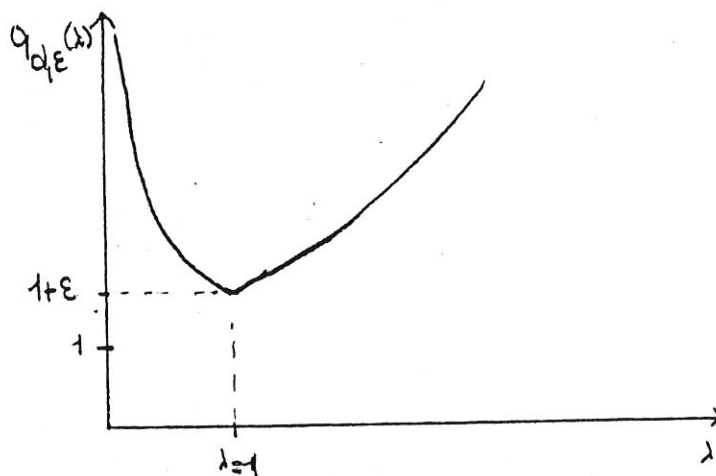
Comme  $q_d(\lambda) = f'(x_0; d) \quad \forall \lambda \leq 1, \quad a_d^* = 1$ .

Nous sommes donc dans le cas où  $0 < a_d^* < +\infty$

Le calcul de  $f''_\varepsilon(x_0; d, d)$  pour  $\varepsilon$  petit nécessite celui de  $f'_\varepsilon(x_0; d)$  et de  $\Lambda_d(\varepsilon)$ .

$$q_{d,\varepsilon}(\lambda) = \begin{cases} 1 + \frac{\varepsilon}{\lambda} & \text{si } \lambda \leq 1 \\ \lambda + \frac{\varepsilon}{\lambda} & \text{si } \lambda > 1 \end{cases}$$

Faisons  $\varepsilon < 1$  et traçons le graphe de  $q_{d,\varepsilon}(\lambda)$



La fonction  $q_{d,\varepsilon}(\lambda)$  atteint sa valeur minimale  $1+\varepsilon$  en  $\lambda=1$

$$\text{Donc } f'_\varepsilon(x_0; d) = 1+\varepsilon$$

$$\Lambda_{d,\varepsilon}(x_0) = \{1\}$$

$$\text{Des lors } f''_\varepsilon(x_0; d, d) = \frac{f'_\varepsilon(x_0; d) - f'(x_0; d)}{\tau_\varepsilon} = \varepsilon$$

$$\text{et } \lim_{\varepsilon \rightarrow 0} f''_\varepsilon(x_0; d, d) = 0.$$

Cet exemple nous permet de résumer les caractéristiques principales dégagées lors de l'étude théorique, dans le cas où  $0 < \alpha_d < +\infty$ .

- comme  $\alpha_d^* = 1$ ,  $f_d$  est affine sur  $[0, 1]$
- $f'_\varepsilon(x_0; d)$  est linéaire en  $\varepsilon$  pour  $\varepsilon$  petit (et vaut  $1 + \varepsilon$ )
- $f''_\varepsilon(x_0; d, d) = \frac{\varepsilon}{(\alpha_d^*)^2} + o(\varepsilon)$  et ici  $o(\varepsilon) = 0$

3)  $\alpha_d^* = 0$

Dans ce cas,  $f''_\varepsilon(x_0; d, d)$  n'est pas nécessairement borné quand  $\varepsilon \rightarrow 0$ , même pour  $d = d$ .

en effet, on a par la proposition 4.1.1:

$$f'_\varepsilon(x_0; d) - f'(x_0; d) \geq \frac{\varepsilon}{\lambda_\varepsilon} \quad \text{ou} \quad \lim_{\varepsilon \rightarrow 0} \frac{\varepsilon}{\lambda_\varepsilon} = 0$$

$$\text{Dès lors } f''_\varepsilon(x_0; d, d) \geq \frac{\varepsilon}{(\lambda_\varepsilon)^2} \geq \frac{\varepsilon}{(\lambda_\varepsilon)^2} \quad (4.2.8)$$

Quand  $\lim_{\varepsilon \rightarrow 0} \frac{\varepsilon}{(\lambda_\varepsilon)^2} = +\infty$ ,  $f''_\varepsilon(x_0; d, d)$  n'est pas bornée.

C'est le cas de  $|x|^\alpha$  de  $\mathbb{R}$  dans  $\mathbb{R}$  (exemple 11). Nous avons

$$\lim_{\varepsilon \rightarrow 0} \frac{\varepsilon}{(\lambda_\varepsilon)^2} = \varepsilon^{\frac{\alpha-3}{2}} (\alpha-1)^{2/d} = +\infty \quad \text{car } \alpha \in ]1, 2[$$

et dès lors  $\lim_{\varepsilon \rightarrow 0^+} f''_\varepsilon(x_0; d, d) = +\infty$ .

De plus, un développement de  $f''_\varepsilon(x_0; d, d)$  tel que (4.3.7) est impossible puisque l'existence dans  $\mathbb{R}^+$  de



$\lim_{\varepsilon \rightarrow 0^+} \left[ f''_{\varepsilon}(x_0; d, d) - \frac{\varepsilon}{(\lambda_{\varepsilon})^2} \right]$  n'est en général pas assurée.

Pour nous en convaincre, calculons cette différence dans ce même exemple.

$$f''_{\varepsilon}(x_0; d, d) - \frac{\varepsilon}{(\lambda_{\varepsilon})^2} = (d-1)^{\frac{d+2}{2}} \varepsilon^{\frac{d-1}{2}}$$

et la limite quand  $\varepsilon$  tend vers 0, est infinie.

La valeur assurant que cette différence est bornée est  $d=2$ .

C'est précisément la valeur minimale telle que  $f(x) = |x|^d$  est "2 fois différentiable". Dans la deuxième partie, nous précisons le sens et donnons la double différentiabilité de la fonction  $f$ .

B.  $f''_{\varepsilon}(x_0; d, d)$  comme approximation de la dérivée de Dini.

---

### 1) Définitions

- $f$  admet une dérivée seconde de Dini, à droite en  $x_0$  dans la direction  $d$  si

$$\lim_{\lambda \rightarrow 0} \frac{1}{\lambda} \left[ f'(x_0 + \lambda d; d) - f'(x_0; d) \right], \text{ notée } D''_n f(x_0; d) \text{ existe finie.} \quad (4.3.9)$$

- $f$  admet une dérivée seconde à droite de la Vallée-Poussin en  $x_0$ , dans la direction  $d$  si :

$$\lim_{\lambda \rightarrow 0} \frac{1}{\lambda} \left[ \frac{f(x_0 + \lambda d) - f(x_0)}{\lambda} - f'(x_0; d) \right], \text{ notée } V''_n f(x_0; d) \text{ existe finie.}$$

### Remarque 1

En règle générale, l'existence de  $D_n'' f(x_0; d)$  assure celle de  $V_n'' f(x_0; d)$  et  $D_n'' f(x_0; d) = 2 V_n'' f(x_0; d)$ . Cependant, pour des fonctions convexes, l'existence de  $D_n'' f(x_0; d)$  et celle de  $V_n'' f(x_0; d)$  sont équivalentes [9].

### Remarque 2

Comme la fonction :  $\lambda \rightarrow f'(x_0 + \lambda d; d)$  est monotone sur  $\mathbb{R}^1$ , la limite (4.3.9) reste inchangée si on remplace  $f'(x_0 + \lambda d; d)$  par des éléments de  $\langle \partial f(x_0 + \lambda d), d \rangle$ . On peut donc réécrire symboliquement :

$$D_n'' f(x_0; d) = \lim_{\lambda \rightarrow 0} \frac{1}{\lambda} \left[ \langle \partial f(x_0 + \lambda d), d \rangle - f'(x_0; d) \right].$$

Pour plus de détails, nous vous renvoyons à l'annexe 5.

### Remarque 3

Si l'une des limites ci-dessus existe, on dira que  $f$  admet une dérivée seconde en  $x_0$  dans la direction  $d$ .

### 2) Lien avec la différentiabilité au sens de Mignot.

D'après Rockafellar ([10], p 887) ou Mignot ([11] §12), la multi-application  $\partial f(\cdot)$  est différentiable en  $x_0$  si  $f$  est différentiable en  $x_0$  et si il existe une application linéaire notée  $D^2 f(x_0)$

telle que  $\|\partial f(x) - \nabla f(x_0) - D^2 f(x_0)(x-x_0)\| = o(\|x-x_0\|)$

Autrement dit :

$$\forall \eta > 0 \exists \delta > 0 \quad \forall x \text{ t.q. } \|x-x_0\| \leq \delta \quad \forall x^* \in \partial f(x)$$

$$\|x^* - \nabla f(x_0) - D^2 f(x_0)(x-x_0)\| \leq \eta \|x-x_0\|$$

$D^2 f(x_0)$  est appelée dérivée de  $\partial f$  en  $x_0$ .

Mignot a prouvé que :

- $\partial f$  est différentiable presque partout sur  $\mathbb{R}^n$
- Si  $\partial f$  est différentiable en  $x_0$  alors  $D^2 f(x_0)$  est symétrique et semi-définie positive.
- $\partial f$  est différentiable en  $x_0$  ssi il existe une matrice symétrique et définie positive  $H^2 f(x_0)$  telle que
 
$$f(x) = f(x_0) + \langle \nabla f(x_0), x-x_0 \rangle + \frac{1}{2} \langle H^2 f(x_0)(x-x_0), x-x_0 \rangle + o(\|x-x_0\|^2)$$
 Ceci est le développement du second ordre de  $f$  au sens d'Alexandroff.

Le lien avec la différentiabilité au sens de Dini est clarifié par le résultat suivant

Si  $f$  est différentiable dans un voisinage de  $x_0$   
 Alors la dérivée de Dini existe ssi  $f$  admet un développement du second ordre au sens d'Alexandroff

Les démonstrations de ces résultats se trouvent dans [4], p. 18-23

### 3) Résultats fondamentaux sur la convergence de $f''_\varepsilon(x_0; d, d)$

a) Examinons d'abord le cas diagonal.

#### Théorème 4.4

Si  $f$  admet une dérivée seconde en  $x_0$  dans la direction  $d$

$$\text{Alors i) } \lim_{\varepsilon \rightarrow 0} f''_\varepsilon(x_0; d, d) = D''_n f(x_0; d)$$

$$\text{ii) } \lim_{\varepsilon \rightarrow 0} \frac{\varepsilon}{(\lambda_\varepsilon)^2} = \frac{1}{2} D''_n f(x_0; d) = V''_n f(x_0; d)$$

$$\forall \lambda_\varepsilon \in \Lambda_d(\varepsilon)$$

Un exemple typique de fonctions  $f$  vérifiant les hypothèses de ce théorème est celui des fonctions du type maximum discret ([12] chap III), c.-à-d.:

$$f(x) = \max \{ f_i(x) \mid i=1, \dots, m \}$$

où les  $f_i$  sont convexes et de classe  $C^2$ .

Pour de telles fonctions,  $D''_n f(x_0; d)$  existe en tout point  $x_0$  et dans toute direction  $d$  et s'exprime à l'aide des formes quadratiques  $\langle \nabla^2 f_i(x_0) d, d \rangle$  associées aux dérivées secondes  $\nabla^2 f_i(x_0)$  des  $f_i$  en  $x_0$ .



### Preuve du théorème

Etudions d'abord le cas où  $a_d^* > 0$

i) Nous avons alors que  $f''_\varepsilon(x_0; d, d) = 0 \quad \forall x_0, \forall d \quad (4.3.6)$

et  $V''_n f(x_0; d) = 0$  (et donc  $D''_n f(x_0; d) = 0$ )

en effet  $V''_n f(x_0; d) = \lim_{\lambda \rightarrow 0} \frac{1}{\lambda} \left[ \frac{f(x_0 + \lambda d) - f(x_0)}{\lambda} - f'(x_0; d) \right]$

or pour  $\lambda$  petit,  $\lambda \in ]0, a_d^*[$  et donc

$\frac{f(x_0 + \lambda d) - f(x_0)}{\lambda} = f'(x_0; d)$  et le crochet s'annule.

En conclusion:

$$\lim_{\varepsilon \rightarrow 0} f''_\varepsilon(x_0; d, d) = D''_n f(x_0; d) = 0$$

ii) la proposition 4.1.1. nous assure que  $\lim_{\varepsilon \rightarrow 0} \frac{\varepsilon}{\lambda_\varepsilon} = 0$  car  $\lambda_\varepsilon \in \Lambda_d(\varepsilon)$

Or  $\lim_{\varepsilon \rightarrow 0} \frac{1}{\lambda_\varepsilon} = \frac{1}{a_d^*}$  qui est un nombre fini.

Des lors  $\lim_{\varepsilon \rightarrow 0} \frac{\varepsilon}{(\lambda_\varepsilon)^2} = 0 = D''_n f(x_0; d) = V''_n f(x_0; d)$ .

Envisageons ensuite le cas où  $a_d^* = 0$

i) La proposition 3.2.4 affirme que

$$f'_\varepsilon(x_0; d) \in \langle \partial f(x_0 + \bar{\lambda}_d(\varepsilon) d), d \rangle$$

et donc

$$f''_\varepsilon(x_0; d, d) \in \frac{\langle \partial f(x_0 + \bar{\lambda}_d(\varepsilon) d), d \rangle - f'(x_0; d)}{\bar{\lambda}_d(\varepsilon)} \quad (4.3.10)$$

D'après le corollaire 1 du théorème 4.1, nous savons que,

quand  $a_d^* = 0$ ,  $\lim_{\varepsilon \rightarrow 0} \bar{\lambda}_d(\varepsilon) = 0$ .

Dès lors, prendre la limite quand  $\varepsilon$  tend vers 0 revient à prendre la limite quand  $\lambda_d(\varepsilon)$  tend vers 0 dans (4.3.10)

Puisque

$$D_n'' f(x_0; d) = \lim_{\lambda \rightarrow 0} \frac{1}{\lambda} [\partial f(x_0 + \lambda d), d] - f'(x_0; d)]$$

on tire de (4.3.10) la première assertion du théorème:

$$\lim_{\varepsilon \rightarrow 0} f''_{\varepsilon}(x_0; d, d) = D_n'' f(x_0; d)$$

ii) Pour prouver la deuxième assertion, calculons la différence.

$$\begin{aligned} f''_{\varepsilon}(x_0; d, d) - \frac{\varepsilon}{[\lambda_d(\varepsilon)]^2} &= \frac{f'_{\varepsilon}(x_0; d) - f'(x_0; d)}{\lambda_d(\varepsilon)} - \frac{\varepsilon}{[\lambda_d(\varepsilon)]^2} \\ &= \frac{1}{\lambda_d(\varepsilon)} \left[ \frac{f(x_0 + \lambda_d(\varepsilon)d) - f(x_0)}{\lambda_d(\varepsilon)} - f'(x_0; d) \right] \end{aligned}$$

Le membre de droite de cette égalité converge, par définition, vers  $V_n'' f(x_0; d)$ , qui vaut  $\frac{1}{2} D_n'' f(x_0; d)$  tandis que  $f''_{\varepsilon}(x_0; d, d)$  converge vers  $D_n'' f(x_0; d)$  par (i).

$$\text{D'où } \lim_{\varepsilon \rightarrow 0} \frac{\varepsilon}{(\lambda_d(\varepsilon))^2} = V_n'' f(x_0; d) = \frac{1}{2} D_n'' f(x_0; d).$$

Enfin, pour obtenir ce résultat avec  $\lambda_d(\varepsilon)$  quelconque dans  $\Lambda_d(\varepsilon)$ , considérons  $\varepsilon' < \varepsilon$

$$\text{Dans ce cas } \lambda_d(\varepsilon') \leq \lambda_d(\varepsilon) \leq \lambda_d(\varepsilon) \leq \bar{\lambda}_d(\varepsilon).$$

$$\text{Ceci implique } \frac{\varepsilon}{(\bar{\lambda}_d(\varepsilon))^2} \leq \frac{\varepsilon}{(\lambda_d(\varepsilon))^2} \leq \frac{\varepsilon}{(\lambda_d(\varepsilon'))^2}.$$

$$\text{D'où } V_n'' f(x_0; d) = \lim_{\varepsilon \rightarrow 0} \frac{\varepsilon}{(\bar{\lambda}_d(\varepsilon))^2} = \lim_{\varepsilon \rightarrow 0} \frac{\varepsilon}{(\lambda_d(\varepsilon'))^2} = \lim_{\varepsilon \rightarrow 0} \frac{\varepsilon}{(\lambda_d(\varepsilon))^2}.$$

Du théorème 4.4, il découle une conséquence importante.

Corollaire 1

Si  $D_n'' f(x_0; d)$  est non nul

$$\text{Alors } \frac{f'_\varepsilon(x_0; d) - f'(x_0; d)}{[2\varepsilon D_n'' f(x_0; d)]^{1/2}} \rightarrow 1 \text{ quand } \varepsilon \rightarrow 0$$

Preuve

$$\begin{aligned} & \lim_{\varepsilon \rightarrow 0} \frac{(f'_\varepsilon(x_0; d) - f'(x_0; d))^2 \lambda_\varepsilon^2}{2\varepsilon (D_n'' f(x_0; d)) \cdot \lambda_\varepsilon^2} \\ &= \frac{1}{D_n'' f(x_0; d)} \cdot \lim_{\varepsilon \rightarrow 0} \frac{(f'_\varepsilon(x_0; d) - f'(x_0; d))^2}{\lambda_\varepsilon^2} \cdot \lim_{\varepsilon \rightarrow 0} \frac{\lambda_\varepsilon^2}{2\varepsilon} \\ &= (D_n'' f(x_0; d))^2 \cdot \frac{1}{(D_n'' f(x_0; d))^2} = 1 \end{aligned}$$

Dans un tel cas, on peut donc développer  $f'_\varepsilon(x_0; d)$  :

$$f'_\varepsilon(x_0; d) = f'(x_0; d) + [2\varepsilon D_n'' f(x_0; d)]^{1/2} + o(\varepsilon^{1/2})$$

et tirer que

$$f'_\varepsilon(x_0; d) - f'(x_0; d) = o(\varepsilon^{1/2})$$

Une illustration typique de cette relation est l'exemple de la fonction quadratique définie sur  $\mathbb{R}^n$  (exemple 6).

Soit  $f$  définie comme

$$f(x) = \frac{1}{2} \langle Ax, x \rangle + \langle b, x \rangle + c$$

où  $A$  est symétrique, définie positive,

$b$  est un vecteur de  $\mathbb{R}^n$ ,  $c$  un nombre réel.

Pour un point  $x_0$  et une direction  $d$ , nous avons calculé que:

$$f'_\varepsilon(x_0; d) = f'(x_0; d) + (2\varepsilon \langle Ad, d \rangle)^{1/2}$$

$$\text{Or } D''_n f(x_0; d) = \langle Ad, d \rangle$$

Donc  $f'_\varepsilon(x_0; d) = f'(x_0; d) + (2\varepsilon D''_n f(x_0; d))^{1/2}$  et  $o(\varepsilon^{1/2})$  est nul.

b) Examinons ensuite le cas général :  $f''_\varepsilon(x_0; d, \delta)$  comme approximation de  $\langle A^2 f(x_0) d, \delta \rangle$

Rappelons que  $A^2 f(x_0)$  est l'application linéaire définie en  $x_0$  où  $\partial f$  est différentiable et vérifie:

$$\forall \varepsilon > 0 \exists \delta > 0 \quad \forall x \text{ t.q. } \|x - x_0\| \leq \delta, \quad \forall x^* \in \partial f(x_0)$$

$$\|x^* - \nabla f(x_0) - A^2 f(x_0)(x - x_0)\| \leq \varepsilon \|x - x_0\|.$$

Pour tout  $x_0$  dans un ensemble  $E$  dont le complémentaire est de mesure nulle, il existe une telle application linéaire  $A^2 f(x_0)$ .

#### Théorème 4.5

Soit  $x_0 \in E$

$$\begin{aligned} \text{Alors } \lim_{\varepsilon \rightarrow 0} f''_\varepsilon(x_0; d, \delta) &= \lim_{\varepsilon \rightarrow 0} \left[ \limsup_{x \rightarrow x_0} f'_\varepsilon(x; d, \delta) \right] \\ &= \langle A^2 f(x_0) d, \delta \rangle \end{aligned}$$

pour toute direction  $\delta$ .

La démonstration de ce théorème se trouve en ([2], p 38-39)



c) Y'intéressons - nous maintenant aux points  $x_0$  n'appartenant pas à  $E$  c.à.d où  $\partial f$  n'est pas différentiable.

Pour ce, nous nous plaçons uniquement dans le cas diagonal. Nous savons que

$$f''_E(x_0; d, d) = \frac{f'_E(x_0; d) - f'(x_0; d)}{\bar{\lambda}_E} \quad \text{où } \bar{\lambda}_E \in \lambda_d(E)$$

et par la proposition 3.2.4

$$f''_E(x_0; d, d) \leq \frac{f'(x_0 + \bar{\lambda}_E d) - f'(x_0; d)}{\bar{\lambda}_E} \quad (4.3.10)$$

On ne peut espérer une limite finie pour  $f''_E(x_0; d, d)$  en tout  $x_0$ . Pour s'en convaincre, d'exemple 4 de la fonction définie par  $f(x) = |x|^2$  nous donne  $\lim_{\varepsilon \rightarrow 0} f''_E(x; 1, 1) = +\infty$ .

A défaut d'avoir nécessairement une limite, nous ferons des hypothèses supplémentaires sur le comportement de

$f_d : \lambda \rightarrow f(x_0 + \lambda d)$  sur  $\mathbb{R}^+$ . Elles assureront que  $\limsup_{\varepsilon \rightarrow 0} f''_E(x_0; d, d) < +\infty$

Pour cela, énonçons d'abord une définition.

La dérivée directionnelle de  $f$  est point-Lipshitz en  $x_0$

dans la direction  $d$  si

$$\limsup_{\lambda \rightarrow 0^+} [f'(x_0 + \lambda d; d) - f'(x_0; d)] / \lambda < +\infty \quad (L_1)$$

Remarquons que sous cette condition et grâce à (4.3.10),

$$\limsup_{\varepsilon \rightarrow 0^+} f''_E(x_0; d, d) < +\infty$$

La proposition suivante fournit une condition nécessaire et suffisante pour que  $f'(x_0; d)$  soit point-lipshitz en  $x_0$ . Elle est démontrée en ([2], p. 40)

Proposition 4.3.1

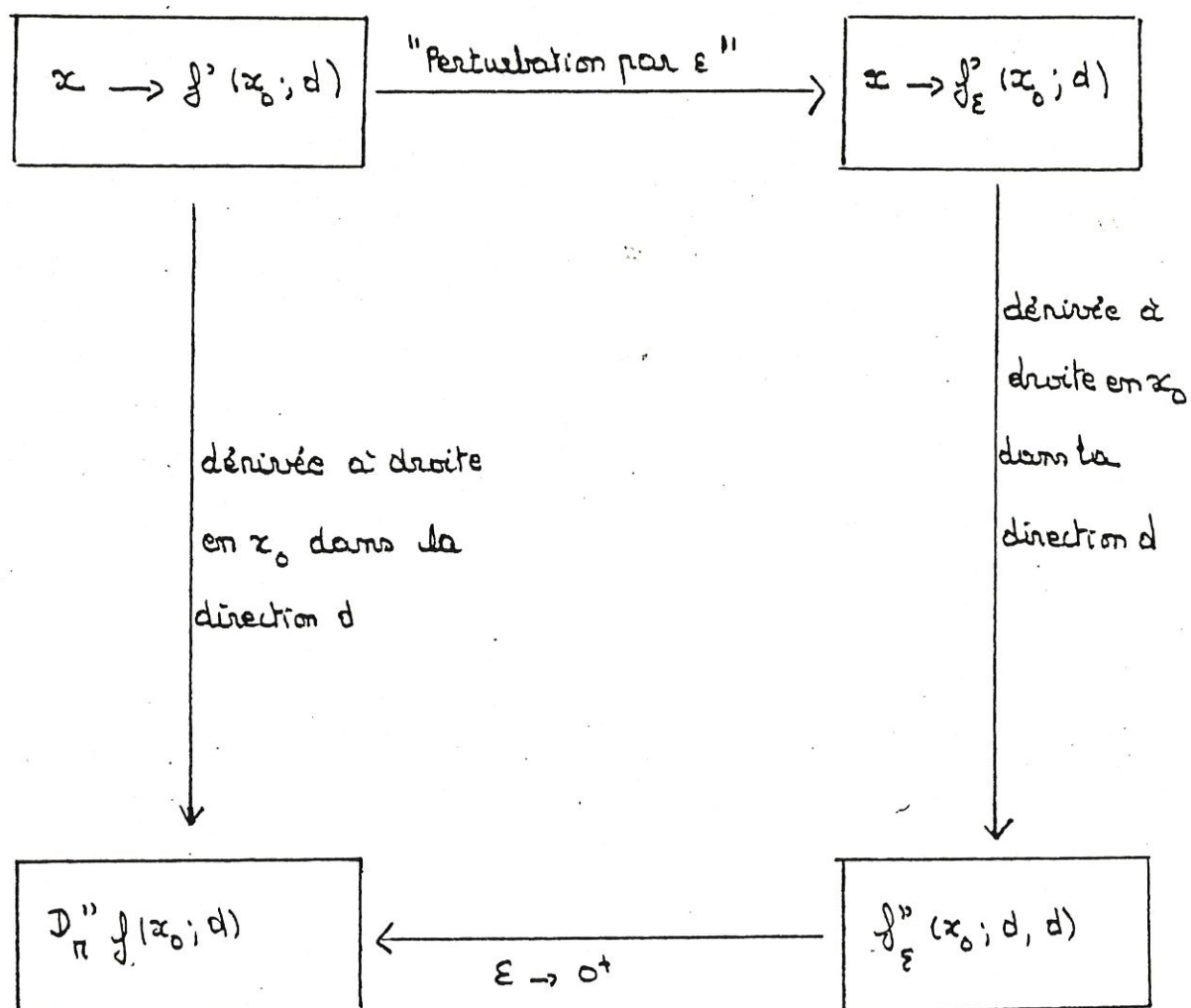
$f'(x_0; d)$  est point-lipshitz en  $x_0$  dans la direction  $d$ ssi

$$\limsup_{\lambda \rightarrow 0^+} \frac{1}{\lambda} \left[ \frac{f(x_0 + \lambda d) - f(x_0)}{\lambda} - f'(x_0; d) \right] < +\infty$$

En fait, la condition  $(L_1)$  est une faible hypothèse sur le comportement de  $f_d$  sur  $\mathbb{R}^+$ . Elle est satisfaite en tout  $x_0$  et pour toute direction  $d$  quand  $f$  est  $\mathcal{C}^{1,1}$ , c.-à-d. quand  $f$  est différentiable et sa dérivée localement lipshitz.

---

Pour résumer l'étude de  $f''_\varepsilon(x_0; d, d)$  comme approximation de la dérivée seconde d'une fonction  $f$ , quand celle-ci est deux fois différentiable en  $x_0$ , examinons le schéma suivant:



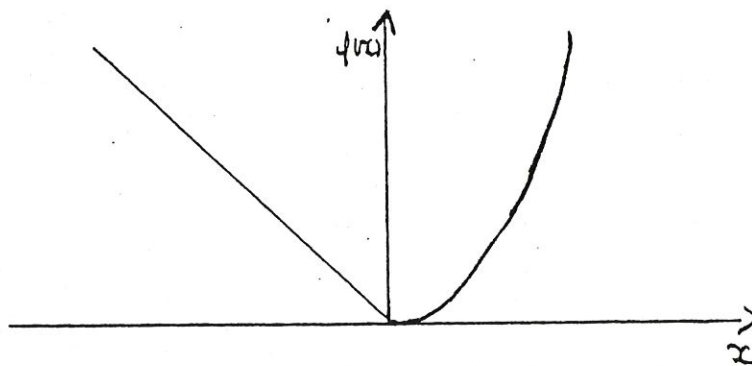
Pour concrétiser les résultats de ce troisième paragraphe, considérons l'exemple suivant.

### exemple 13

Soit  $f: \mathbb{R} \longrightarrow \mathbb{R}$

$$x \longmapsto \begin{cases} \frac{1}{3} x^3 & \text{si } x \geq 0 \\ -x & \text{si } x \leq 0 \end{cases}$$

Faisons  $x_0 = 0$ ,  $d = 1$



On voit sur le graphe que  $\partial f(0) = [-1, 0]$

et donc  $f'(0; 1) = 0$

Vérifions que  $\lim_{\varepsilon \rightarrow 0} f''_{\varepsilon}(x_0; d, d) = D''_{\pi} f(x_0; d)$

Pour cela, calculons successivement  $f'_{\varepsilon}(x_0; d)$  et  $\lambda_d(\varepsilon)$ ,  $f''_{\varepsilon}(x_0; d, d)$  et  $D''_{\pi} f(x_0; d)$ .

$$\bullet \quad \varphi_{d, \varepsilon}(\lambda) = \frac{1}{3} \lambda^3 + \frac{\varepsilon}{\lambda}$$

Cette fonction atteint sa valeur minimale  $\frac{4}{3} \varepsilon^{3/4}$  en

$\lambda = \varepsilon^{1/4}$ . Dès lors

$$\lambda_d(\varepsilon) = \{ \varepsilon^{1/4} \}$$

$$f'_{\varepsilon}(x_0; d) = \frac{4}{3} \varepsilon^{3/4}$$



$$\cdot f''_{\varepsilon}(x_0; d, d) = \frac{f'_{\varepsilon}(x_0; d) - f'(x_0; d)}{\lambda_{\varepsilon}} = \frac{4}{3} \varepsilon^{1/2}$$

$$\cdot D''_n f(x_0; d) = \lim_{\lambda \rightarrow 0} \frac{1}{\lambda} [f'(x_0 + \lambda d; d) - f'(x_0; d)]$$

$$\begin{aligned} \text{... où } f'(x_0 + \lambda d; d) &= \nabla f(x_0 + \lambda d) \cdot d = \frac{4}{3} \lambda^3 \\ &= \lim_{\lambda \rightarrow 0} \frac{4}{3} \lambda^2 = 0 \end{aligned}$$

$$\text{Nous avons donc } \lim_{\varepsilon \rightarrow 0} f''_{\varepsilon}(x_0; d, d) = D''_n f(x_0; d) = 0$$

Remarquons que  $\partial f(\cdot)$  n'est pas différentiable en  $x_0 = 0$  puisque  $\partial f(x_0)$  n'est pas réduit à un singleton. Le point  $x_0$  ne vérifie donc pas l'hypothèse du théorème 4.5. Nous ne pouvons dès lors rien dire quant à la limite de  $f''_{\varepsilon}(x_0; d, d)$ .

Ce chapitre, axé sur le comportement limite de  $f''_{\varepsilon}(x_0; d, d)$  en fonction de  $\varepsilon$ , a été axé en trois parties. La première a été consacrée à l'étude de l'ensemble  $\Lambda_d(\varepsilon)$ . Nous y avons prouvé que la suite des intervalles  $\Lambda_d(\varepsilon)$  converge vers le singleton  $\{a_d^*\}$  quand  $\varepsilon$  tend vers zéro.

Dans la deuxième partie, nous avons étudié le comportement qualitatif de  $f'_{\varepsilon}(x_0; d)$ , encore notée  $\Theta_d(\varepsilon)$ . Nous avons déduit que  $M_d(\varepsilon)$  est un singleton pour presque tous les  $\varepsilon$  et que

$(\Theta_d)'_+(0) = \frac{1}{a_d^*}$ . Ceci nous a suggéré un développement du premier ordre pour  $f'_\varepsilon(x_0; d)$  dans le cas où  $0 < a_d^* < +\infty$ . Enfin, dans la troisième partie, nous avons défini la dérivée de Diré dans la direction  $d$  et prouvé que, en un point  $x_0$  où celle-ci existe, elle est égale à la limite de  $f'_\varepsilon(x_0; d, d)$  quand  $\varepsilon$  tend vers 0. D'autre part, si  $f$  admet un développement au sens d'Alexandrov en  $x_0$ ,  $f'_\varepsilon(x_0; d, \delta)$  converge vers  $(A^2 f(x_0) d, \delta)$  où  $A^2 f(x_0)$  tient lieu de Hessian dans le développement. Cette approximation à  $\varepsilon$ -près est donc une bonne généralisation du concept de dérivée seconde. Afin de compléter l'étude de  $f'_\varepsilon(x_0; d, \delta)$ , nous nous proposons maintenant de l'examiner comme fonction de  $x$  et de  $(d, \delta)$ .

## CHAPITRE V.

Etude de  $f_\epsilon''(x; d, \delta)$  comme  
fonction de  $x$  et  $(d, \delta)$ .

Dans ce chapitre, nous relierons les derniers résultats récoltés jusqu'à présent sur  $f'_\varepsilon(x; d, \delta)$ . Signalons que la liste est loin d'être exhaustive puisque ces éléments font partie d'un domaine où la recherche est encore totalement ouverte.

## II.1 Semi-continuité de $f'_\varepsilon(x; d, d)$ comme fonction de $x$ .

Rappelons que  $f'_\varepsilon(\cdot; d, d)$  est la fonction qui à  $x$  fait correspondre  $\bar{\mu}_{d, \varepsilon}(x) [f'_\varepsilon(x; d) - f'(x; d)]$ , où  $\bar{\mu}_{d, \varepsilon}(x)$  est le minimum de l'intervalle  $\Pi_{d, \varepsilon}(x)$ .

Afin d'étudier la continuité de  $f'_\varepsilon(x; d, d)$  en fonction de  $x$ , examinons chacun des facteurs qui la constitue. D'une part, nous savons que la multiplication  $x \mapsto \Pi_{d, \varepsilon}(x)$  est semi-continue supérieurement en chaque  $x_0$  et bornée dans un voisinage de  $x_0$ . Dès lors, nous montrons que l'application  $x \mapsto \bar{\mu}_{d, \varepsilon}(x)$  est semi-continue inférieurement i.e.

Pour toute suite  $(x_n)_{n \in \mathbb{N}}$  de  $\mathbb{R}^n$  telle que  $(x_n) \rightarrow x$

$$\lim_{n \rightarrow \infty} \bar{\mu}_{d, \varepsilon}(x_n) \geq \bar{\mu}_{d, \varepsilon}(x).$$



En effet :

Notons la plus petite valeur d'adhérence de la suite  $(\bar{\mu}_{d,\varepsilon}(x_n))_n$ ,  
 $\lim_{n \rightarrow \infty} \bar{\mu}_{d,\varepsilon}(x_n) = a$ . On peut donc extraire une sous-suite  
 encore appelée  $(\bar{\mu}_{d,\varepsilon}(x_n))_n$  qui converge vers  $a$  où chaque  $\bar{\mu}_{d,\varepsilon}(x_n)$   
 appartient à  $\Pi_{d,\varepsilon}(x_n)$ . Le caractère fermé de  $\Pi_{d,\varepsilon}(\cdot)$  nous  
 garantit que  $a \in \Pi_{d,\varepsilon}(x)$ . Dès lors, par définition de  $\bar{\mu}_{d,\varepsilon}(x)$ ,  
 nous avons :  $a \geq \bar{\mu}_{d,\varepsilon}(x)$

□.

D'autre part, nous savons que

- la fonction :  $x \mapsto f'(x; d)$  est s.-c. supérieurement (I. 4.16)  
 et donc  $x \mapsto -f'(x; d)$  est s.-c. inférieurement
- la fonction :  $x \mapsto f'_\varepsilon(x; d)$  est L.L. et donc continue.

En conclusion,  $f''_\varepsilon(x; d, d)$  en tant que produit de deux  
 fonctions positives et s.e.i., est elle-même une fonction positive et s.e.i.

Remarque.

Si  $\Pi_{d,\varepsilon}(x_0)$  est un singleton et si  $f$  est différentiable en  $x_0$ ,  
 alors  $\bar{\mu}_{d,\varepsilon}(\cdot)$  et  $f'(\cdot; d)$  (où  $f'(\cdot; d) = Df(\cdot) \cdot d$ ) sont deux  
 fonctions continues. Dès lors,  $f''_\varepsilon(\cdot; d, d)$  l'est aussi.

V.2. Comportement de  $f''_{\varepsilon}(x_0; d, \delta)$  comme fonction de  $(d, \delta)$ .

A.  $f''_{\varepsilon}(x_0; d, \delta)$  est positivement homogène en  $(d, \delta)$ .

Pour étudier  $f''_{\varepsilon}$  en  $(d, \delta)$ , examinons d'abord le comportement de  $\Pi_{d, \varepsilon}(x_0)$  comme multiplication de  $d$ . A ce propos, nous avons la relation suivante :

Lemme 1.

$$\text{Pour tout } \alpha > 0 : \quad \Pi_{\alpha d, \varepsilon}(x_0) = \alpha \Pi_{d, \varepsilon}(x_0).$$

Preuve.

Pour  $x_0$  fixé, on sait que  $\partial \Theta_d(\varepsilon) = \Pi_{d, \varepsilon}(x_0)$  pour tout  $\varepsilon > 0$

où  $\Theta_d : \varepsilon \mapsto f'_{\varepsilon}(x_0; d)$ . ( )

Or  $f'_{\varepsilon}(x_0; \alpha d) = \alpha f'_{\varepsilon}(x_0; d)$  pour tout  $\alpha > 0$  (I.4.13)

ou bien  $\Theta_{\alpha d}(\varepsilon) = \alpha \Theta_d(\varepsilon)$  pour tout  $\alpha > 0$ .

et donc  $\partial \Theta_{\alpha d}(\varepsilon) = \alpha \partial \Theta_d(\varepsilon)$  pour tout  $\alpha > 0$ .

□

Ceci nous permet de prouver que  $f''_{\varepsilon}(x_0; d, \delta)$  est positivement homogène de degré 2.

Proposition 5.2.1.

$$f''_{\varepsilon}(x_0; \alpha d, \beta \delta) = f''_{\varepsilon}(x_0; \beta d, \alpha \delta) = \alpha \beta f''_{\varepsilon}(x_0; d, \delta) \quad \forall \alpha, \beta > 0$$

En effet : pour  $\alpha, \beta > 0$ , quelconques :

$$\begin{aligned} f''_{\varepsilon}(x_0; \alpha d, \beta \delta) &= \min_{\mu \in \Pi_{\alpha d, \varepsilon}(x_0)} \max_{x^* \in \partial_{\varepsilon} f(x_0)_{\alpha d}} \mu [\langle x^*, \beta \delta \rangle - f'(x_0; \beta \delta)] \\ &= \beta \min_{\mu \in \Pi_{\alpha d, \varepsilon}(x_0)} \max_{x^* \in \partial_{\varepsilon} f(x_0)_{\alpha d}} \mu [\langle x^*, \delta \rangle - f'(x_0; \delta)] \end{aligned}$$

car  $f'(x_0; \delta)$  est positivement homogène en  $\delta$ .

Or  $\Pi_{\alpha d, \varepsilon}(x_0) = \alpha \Pi_{d, \varepsilon}(x_0)$  par le lemme 1.

$$\begin{aligned} \text{or } \partial_{\varepsilon} f(x_0)_{\alpha d} &= \{x^* \in \partial_{\varepsilon} f(x_0) \mid \langle x^*, \alpha d \rangle = f'_{\varepsilon}(x_0; \alpha d)\} \\ &= \partial_{\varepsilon} f(x_0)_d \end{aligned}$$

Dès lors, en posant  $\mu = \alpha \mu'$  on obtient :

$$\begin{aligned} f''_{\varepsilon}(x_0; \alpha d, \beta \delta) &= \beta \alpha \min_{\mu' \in \Pi_{d, \varepsilon}(x_0)} \max_{x^* \in \partial_{\varepsilon} f(x_0)_d} \mu' [\langle x^*, \delta \rangle - f'(x_0; \delta)] \\ &= \beta \alpha f''_{\varepsilon}(x_0; d, \delta) \end{aligned}$$

□

B. Comportement de  $f''_{\varepsilon}(x_0; d, d)$  quand  $\|d\| \rightarrow \infty$ .

Le comportement peut être décrit par la proposition suivante :

Proposition 5.2.2.

Soit  $w = d/\|d\|$ . Nous avons

$$\lim_{\|d\| \rightarrow +\infty} f''_{\varepsilon}(x_0; d, d) = +\infty \quad \text{pour tout } 0 < \varepsilon < \varepsilon_w^*$$

Preuve.

L'inégalité déjà connue sous la forme :

$$f''_{\varepsilon}(x_0; d, d) \geq \varepsilon [\bar{\mu}_d(\varepsilon)]^2 \quad (\text{IV. 3.2})$$

est améliorée par le lemme 1 en

$$f''_{\varepsilon}(x_0; d, d) \geq \varepsilon [\bar{\mu}_{\omega}(\varepsilon)]^2 \|d\|^2 \quad \text{où } \omega = \frac{d}{\|d\|}.$$

Par conséquent,  $\lim_{\|d\| \rightarrow +\infty} f''_{\varepsilon}(x_0; d, d) = +\infty$  si  $\bar{\mu}_{\omega}(\varepsilon)$  ne

tend pas vers 0.

Or puisque  $\varepsilon_{\omega}^* > 0$ , nous avons :

$$\Pi_{\omega, \varepsilon}(x_0) \subset \mathbb{R}_+^+ \text{ pour chaque } 0 < \varepsilon < \varepsilon_{\omega}^*.$$

Dès lors :  $\lim_{\|d\| \rightarrow \infty} f''_{\varepsilon}(x_0; d, d) = +\infty$  pour tout  $0 < \varepsilon < \varepsilon_{\omega}^*$ .

□

C. Convexité de  $f''_{\varepsilon}(x_0; d, d)$  en fonction de  $d$ .

La question de la convexité de  $f''_{\varepsilon}(x_0; d, d)$  a été élucidée par Lemaréchal (1980). Même dans un cas simple comme celui des fonctions polyédrales, il a montré que :

Pour un point  $x_0$  fixé, où  $f$  n'est pas différentiable,  $f''_{\varepsilon}(x_0; d, d)$  n'est pas convexe en  $d$ .

Développons son contre-exemple :



Considérons  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  définie par

$$f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \max \{ x, y, -1+x+y \}.$$

Prenons  $X_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$  et  $\varepsilon \in ]0, 1[$ .

et  $d_\alpha = \begin{pmatrix} \alpha \\ 1-\alpha \end{pmatrix}$  avec  $\alpha \in [0, 1]$ .

Nous calculerons donc  $f''_\varepsilon(X_0; d_\alpha, d_\alpha)$  pour une direction  $d_\alpha$  combinaison convexe de  $d_0$  et  $d_1$ , i.e.  $d_\alpha = \alpha d_1 + (1-\alpha) d_0$ .

Nous vérifierons que l'on n'a pas :

$$f''_\varepsilon(X_0; d_\alpha, d_\alpha) \leq \alpha f''_\varepsilon(X_0; d_1, d_1) + (1-\alpha) f''_\varepsilon(X_0; d_0, d_0).$$

Pour cela, nous recherchons  $f'(X_0; d_\alpha)$ ,  $f'_\varepsilon(X_0; d_\alpha)$ ,  $\Lambda_{d_\alpha, \varepsilon}(X_0)$  et enfin  $f''_\varepsilon(X_0; d_\alpha, d_\alpha)$  pour  $\alpha \in [0, 1]$ .

1)  $f'(X_0; d_\alpha)$  est la fonction d'appui du sous-différentiel de  $f$  en

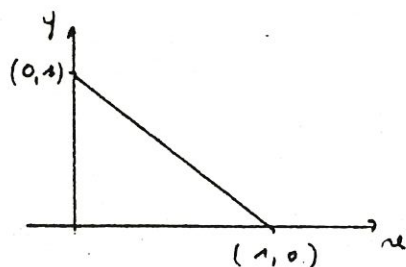
$$\partial f(X_0) = \text{co} \{ \nabla f_i(X_0), i \text{ tq } f(X_0) = f_i(X_0) \}$$

$$= \text{co} \left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$$

$$f'(X_0; d_\alpha) = \max_{x^* \in \partial f(0)} \langle x^*, d \rangle$$

$$= \max_{\begin{pmatrix} x_1^* \\ x_2^* \end{pmatrix} \in \partial f(0)} (x_1^* \alpha + x_2^* (1-\alpha))$$

Représentons  $\partial f(0)$ .



et donc  $f'(x_0; d_\alpha) = \max(\alpha, 1-\alpha)$ .

2) Ensuite, calculons  $\lambda_{d_\alpha, \varepsilon}(x_0)$  et  $f'_\varepsilon(x_0; d_\alpha)$ .

Pour cela, remarquons que le minimum de

$$q_{d_\alpha, \varepsilon}(\lambda) \equiv \frac{f(x_0 + \lambda d_\alpha) - f(x_0) + \varepsilon}{\lambda} \quad \text{est atteint en } \lambda$$

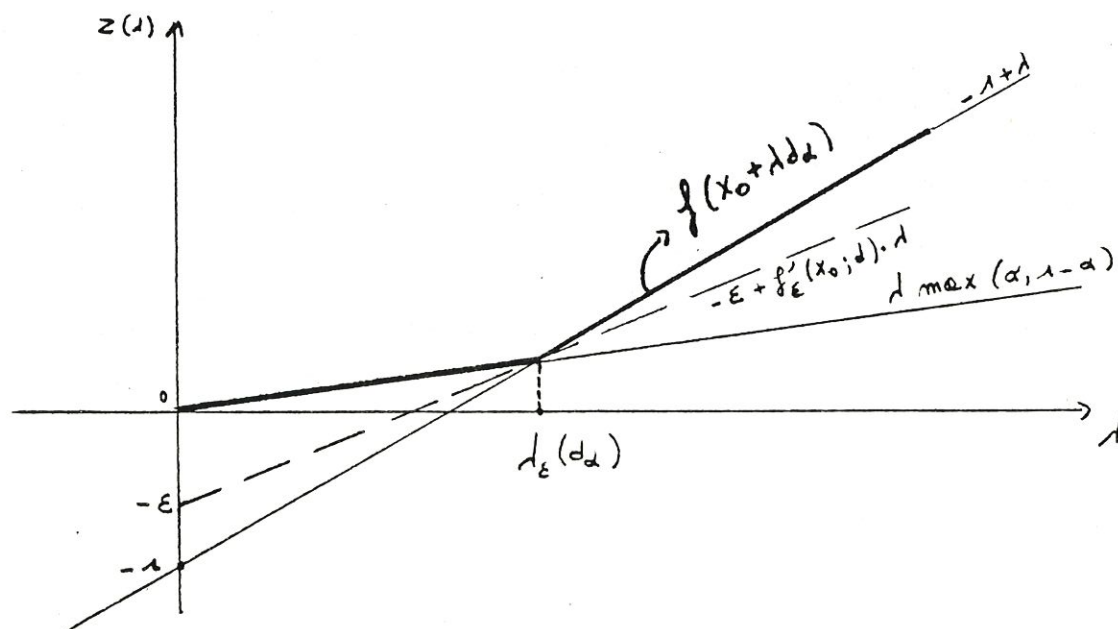
$$\text{si ce } \lambda \text{ vérifie } \frac{f(x_0 + \lambda d_\alpha) - f(x_0) + \varepsilon}{\lambda} = f'_\varepsilon(x_0; d_\alpha).$$

$$\text{i.e.} \quad f(x_0 + \lambda d_\alpha) = \lambda f'_\varepsilon(x_0; d_\alpha) + f(x_0) - \varepsilon.$$

Or  $f(x_0) = 0$ , et

$$\begin{aligned} f(x_0 + \lambda d_\alpha) &= \max(\lambda, \lambda(1-\alpha), -1 + \lambda\alpha + \lambda(1-\alpha)) \\ &= \max(\lambda \max(\alpha, 1-\alpha), -1 + \lambda). \end{aligned}$$

• Recherchons ce point  $\lambda_\varepsilon(d_\alpha)$  pour  $\alpha \in ]0, 1[$  graphiquement.



Le point  $\lambda_\varepsilon(d_\alpha)$  est le point d'intersection des deux droites d'équation

$$\begin{cases} z = -1 + \lambda \\ z = \lambda \max(\alpha, 1-\alpha) \end{cases}$$

et donc

$$\Lambda_{d_\alpha, \varepsilon}(x_0) = \{ \lambda_\varepsilon(d_\alpha) \} = \left\{ \frac{1}{1 - \max(\alpha, 1-\alpha)} \right\}$$

$$\begin{aligned} \text{Dès lors, } f'_\varepsilon(x_0; d) &= \frac{\lambda_\varepsilon(d_\alpha) \max(\alpha, 1-\alpha) + \varepsilon}{\lambda_\varepsilon(d_\alpha)} \\ &= f'(x_0; d_\alpha) + \varepsilon [1 - \max(\alpha, 1-\alpha)] \end{aligned}$$

• Pour  $\alpha = 0$  ou  $\alpha = 1$ , on a

$$f(x_0 + \lambda d_\alpha) = 1.$$

Par un même raisonnement, il suit que

$$\Lambda_{d_1, \varepsilon}(x_0) = \Lambda_{d_0, \varepsilon}(x_0) = \emptyset.$$

3). Le calcul de  $f''_\varepsilon(x_0; d_\alpha, d_\alpha)$  est immédiat pour  $\alpha = 0$  et  $\alpha = 1$  puisque dans ces cas,  $\Lambda_{d_\alpha, \varepsilon}(x_0) = \{0\}$ .

$$\text{Dès lors, } f''_\varepsilon(x_0; d_0, d_0) = f''_\varepsilon(x_0; d_1, d_1) = 0.$$

• Pour  $\alpha \in ]0, 1[$ , on a

$$\begin{aligned} f''_\varepsilon(x_0; d_\alpha, d_\alpha) &= \frac{f'_\varepsilon(x_0; d_\alpha) - f'(x_0; d_\alpha)}{\lambda_\varepsilon(d_\alpha)} = \varepsilon [1 - \max(\alpha, 1-\alpha)]^2 \\ &= \varepsilon [\min(\alpha, 1-\alpha)]^2. \end{aligned}$$

$$\text{et donc } f''_\varepsilon(x_0; d_\alpha, d_\alpha) > 0.$$

L'inégalité :

$$f''_\varepsilon(x_0; d_\alpha, d_\alpha) \leq \alpha f''_\varepsilon(x_0; d_1, d_1) + (1-\alpha) f''_\varepsilon(x_0; d_0, d_0)$$

avec  $d_\alpha = \alpha d_1 + (1-\alpha) d_0$ , est donc impossible

Ceci confirme la non convexité de  $f'_\varepsilon(x_0; d, d)$  en fonction de  $d$ .  
 Signalons que le point  $x_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$  est un point où  $f$  n'est pas différentiable. C'est précisément cette raison qui empêche la convexité.

Proposition 5.2.3.

Si  $f$  est une fonction polyédrale, différentiable en  $x_0$ .  
 Alors, il existe  $\underline{\varepsilon} \geq 0$  tel que  
 $f''_\varepsilon(x_0; d, d)$  est convexe en  $d$  pour tout  $0 < \varepsilon < \underline{\varepsilon}$ .

Illustrons cette proposition qui est démontrée en ([4], p. 29-30),  
 par le même exemple, mais considérons maintenant un point  
 $x_0$  où  $f$  est différentiable.

Choisissons  $x_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ,  $d = \begin{pmatrix} \alpha \\ 1-\alpha \end{pmatrix}$ ,  $\alpha \in [0, 1]$   
 et  $0 < \varepsilon < 1$ .

1) Dès lors  $f(x_0) = 1$ , et

$$\nabla f(x_0) = \text{co} \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ et } f'(x_0; d) = \alpha$$

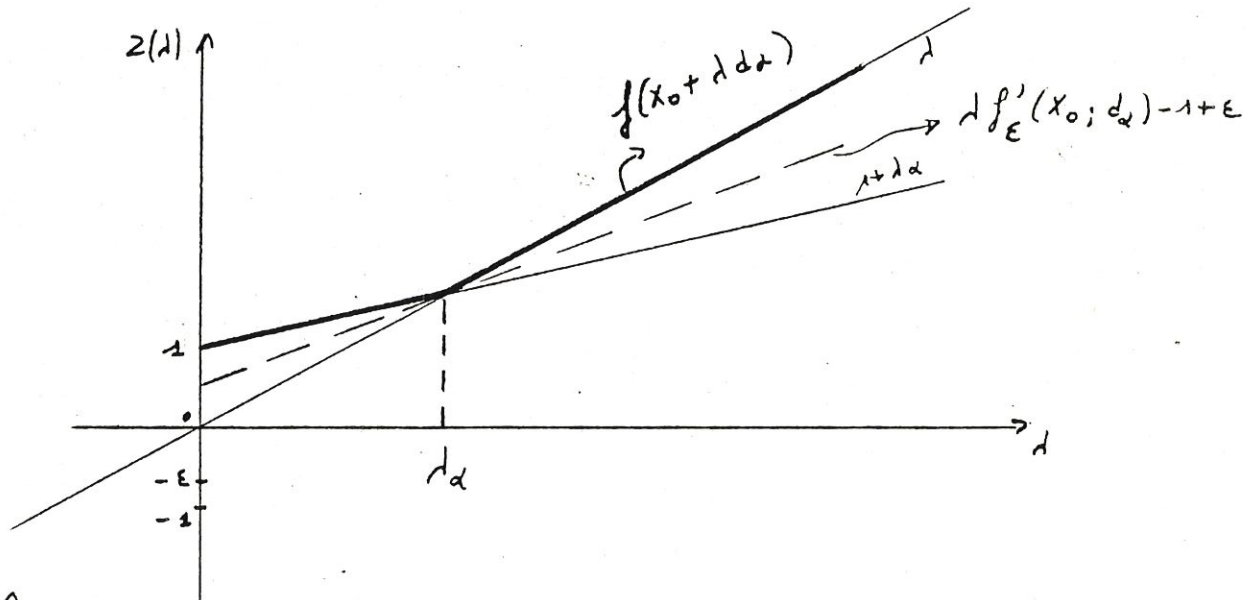
2) Calculons  $\Lambda_{d, \varepsilon}(x_0)$ ,  $f'_\varepsilon(x_0; d, d)$  et donc  $q_{d, \varepsilon}(1)$ .

$$\begin{aligned} q_{d, \varepsilon}(1) &= \frac{f(x_0 + 1d) - f(x_0) + \varepsilon}{1} \\ &= \frac{\max(1, 1 + 1\alpha) - 1 + \varepsilon}{1} \end{aligned}$$

Le  $d_\alpha$  qui minimise  $q_{d, \varepsilon}(1)$  est tel que



$f(x_0 + \lambda d_\alpha) = \max(\lambda_\alpha, 1 + \lambda_\alpha) = \lambda_\alpha f'_\epsilon(x_0; d_\alpha) - 1 + \epsilon$   
 Recherchons graphiquement  $\lambda_\alpha$  pour  $\alpha \in ]0, 1[$ .



Le point  $\lambda_\alpha$  est point d'intersection des deux droites d'équations

$$\begin{cases} z(\lambda) = 1 \\ z(\lambda) = 1 + \lambda\alpha \end{cases}$$

et donc

$$\lambda_{d_\alpha, \epsilon}(x_0) = \{\lambda_\alpha\} = \left\{ \frac{1}{1-\alpha} \right\} \text{ et } f'_\epsilon(x_0; d_\alpha) = \epsilon + \epsilon(1-\alpha).$$

3) Calculons  $f''_\epsilon(x_0; d_\alpha, d_\alpha)$  pour tout  $\alpha \in [0, 1]$ .

• Si  $0 < \alpha < 1$ , on a

$$f''_\epsilon(x_0; d_\alpha, d_\alpha) = \epsilon(1-\alpha)^2, \text{ et}$$

• Si  $\alpha = 0$ , on calcule que

$$f'(x_0; d_0) = 0$$

$$q_{d_0, \epsilon}(\lambda) = \frac{\max(\lambda, 1) - 1 + \epsilon}{\lambda}$$

$$q_{d_0, \varepsilon}(d) = \begin{cases} \varepsilon/d & \text{si } d < 1 \\ 1 + \frac{\varepsilon-1}{d} & \text{si } d > 1 \end{cases}$$

$$\text{et } f'_\varepsilon(x_0; d_0) = \inf q_{d_0, \varepsilon}(d) = \varepsilon$$

$$\Lambda_{d_0, \varepsilon}(x_0) = \{1\}$$

$$\text{Dès lors, } f''_\varepsilon(x_0; d_0, d_0) = \varepsilon.$$

• Si  $\alpha = 1$ , on calcule que

$$f'(x_0; d_1) = 1$$

$$q_{d_1, \varepsilon}(d) = 1 + \varepsilon/d$$

$$\text{et } f'_\varepsilon(x_0; d_1) = 1$$

$$\Lambda_{d_1, \varepsilon} = \emptyset$$

$$\text{Dès lors } f''_\varepsilon(x_0; d_1, d_1) = 0.$$

On vérifie alors pour  $d_\alpha = \alpha d_1 + (1-\alpha) d_0$ , où  $d_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$  et  $d_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ ,  
 $\alpha \in [0, 1]$

$$\text{que } f''_\varepsilon(x_0; d_\alpha, d_\alpha) \leq \alpha f''_\varepsilon(x_0; d_1, d_1) + (1-\alpha) f''_\varepsilon(x_0; d_0, d_0)$$

$$\text{car } \varepsilon(1-\alpha)^2 \leq \alpha \cdot 0 + (1-\alpha) \varepsilon.$$

De plus, ce calcul peut être effectué pour des directions  $d_0$  et  $d_1$  quelconques. Ceci illustre bien la convexité de  $f''_\varepsilon(x_0; d, d)$  en la variable  $d$  quand  $f$  est différentiable en  $x_0$ .

---

D. Continuité de  $f''_{\varepsilon}(x_0; d, d)$  en fonction de  $d$ .

---

La fonction  $d \mapsto f''_{\varepsilon}(x_0; d, d)$  n'est en général pas continue.

Exhibons un contre-exemple.

Soit  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ .

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \max(x^2 + y^2, x^2 + x)$$

Choisissons  $x_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

Calculons d'abord la valeur de la fonction  $f''_{\varepsilon}(x_0; d, d)$  pour une direction  $d_{\alpha}$  dans le premier quadrant "ouvert"

ie pour  $d_{\alpha} = \begin{pmatrix} \alpha \\ \sqrt{1-\alpha^2} \end{pmatrix}$ , avec  $0 < \alpha < 1$ , et ensuite pour

$$d_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ et } d_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Nous vérifierons qu'elle n'est pas continue en  $d_0$ .

1)  $d_{\alpha} = \begin{pmatrix} \alpha \\ \sqrt{1-\alpha^2} \end{pmatrix}$  avec  $0 < \alpha < 1$ .

Calculons

$$a) \quad q_{d_{\alpha}}(t) = \max(t, t\alpha^2 + \alpha)$$

$$\text{dès lors } q_{d_{\alpha}}(t) = \begin{cases} t\alpha^2 + \alpha & \text{si } t \leq \frac{\alpha}{1-\alpha^2} \\ t & \text{sinon.} \end{cases}$$

$$\text{et } f'(x_0; d_\alpha) = \lim_{d \rightarrow 0} q_{d_\alpha}(d) = d$$

$$b) q_{d_\alpha, \varepsilon}(d) = \begin{cases} d\alpha^2 + \alpha + \frac{\varepsilon}{d} & \text{si } d \leq \frac{\alpha}{1-\alpha^2} \\ d + \frac{\varepsilon}{d} & \text{si } d \geq \frac{\alpha}{1-\alpha^2} \end{cases}$$

L'infimum de  $q_{d_\alpha, \varepsilon}(d)$  est atteint pour  $d = \frac{\sqrt{\varepsilon}}{\alpha}$ , si  $\frac{\sqrt{\varepsilon}}{\alpha}$  est inférieur à  $\frac{\alpha}{1-\alpha^2}$ . Ce sera le cas si nous choisissons  $\varepsilon$  suffisamment petit.

$$\text{Dès lors, } f'_\varepsilon(x_0; d_\alpha) = 2\sqrt{\varepsilon}\alpha + \alpha.$$

$$c) f''_\varepsilon(x_0; d_\alpha, d_\alpha) = 2\alpha^2.$$

$$2) \underline{d_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}}$$

Calculons

$$a) q_{d_0}(d) = \frac{1}{d} \max(d^2, 0) = d.$$

$$\text{dès lors } f'(x_0; d_0) = 0.$$

$$b) q_{d_0, \varepsilon}(d) = d + \frac{\varepsilon}{d}$$

L'infimum  $f'_\varepsilon(x_0; d_0)$  vaut  $2\sqrt{\varepsilon}$  et est atteint par  $d = \sqrt{\varepsilon}$ .

$$c) f''_\varepsilon(x_0; d_0, d_0) = 2.$$

$$3) \underline{d_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}}$$

Calculons

$$a) q_{d_1}(d) = \max(d, d+1) = d+1.$$

$$\text{dès lors } f'(x_0; d_1) = 1.$$

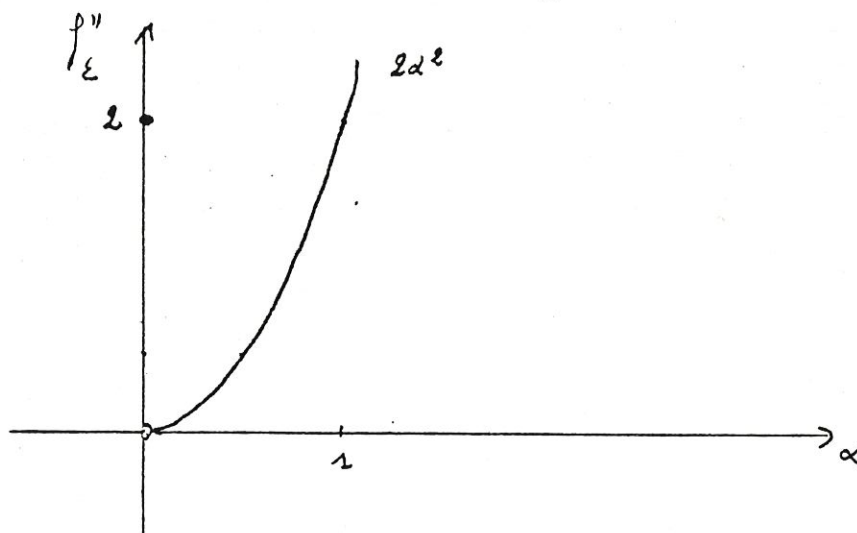


$$b) \varphi_{d_1, \varepsilon}(d) = d + 1 + \frac{\varepsilon}{d}$$

L'infimum  $f'_\varepsilon(x_0; d_1)$  vaut  $2\sqrt{\varepsilon} + 1$  et est atteint par  $d = \sqrt{\varepsilon}$ .

$$c) f''_\varepsilon(x_0; d_1, d_1) = 2.$$

Représentons la fonction  $\alpha \mapsto f''_\varepsilon(x_0; d_\alpha, d_\alpha)$



La fonction  $f''_\varepsilon$  est discontinue en  $\alpha = 0$  et donc  $f''_\varepsilon(x_0; d, d)$  n'est pas continue en  $d_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

### V. 3 Règles de calcul sur $f''_\varepsilon(x_0; d, \delta)$

Comme pour la dérivée directionnelle à  $\varepsilon$  près, il est intéressant d'avoir des expressions de  $(f_1 + f_2)_\varepsilon''$ ,  $(f \circ A)_\varepsilon''$ ,  $(\max_{i=1, \dots, m} f_i)_\varepsilon''$ , etc... en termes de  $(f_i)_\varepsilon''$ .

Pour cela, il est nécessaire de connaître les expressions de  $\Pi_{d,\epsilon}^{f_1+f_2}(x_0)$ ,  $\Pi_{d,\epsilon}^{f \circ A}(x_0)$  etc... en termes de  $\Pi_{d_i,\epsilon_i}^{f_i}(x_i)$ .  
 Ces résultats ne sont pas encore publiés, mais donnons - en toutefois une idée.

Étant donné une fonction convexe  $f$  de  $\mathbb{R}^n$  dans  $\mathbb{R}$  et une application linéaire  $A$  de  $\mathbb{R}^n$  dans  $\mathbb{R}^n$ , nous avons que :

$$(f \circ A)'_{\epsilon}(x_0; d) = f'_{\epsilon}(Ax_0; Ad)$$

$$\Pi_{d,\epsilon}^{f \circ A}(x_0) = \Pi_{Ad,\epsilon}^f(Ax_0) \quad \text{pour tout } x_0, d; \epsilon > 0$$

Dès lors, nous prouvons que

$$(f \circ A)''_{\epsilon}(x_0; d, d) = f''_{\epsilon}(Ax_0; Ad, Ad)$$

En effet :

$$\begin{aligned} (f \circ A)''_{\epsilon}(x_0; d, d) &= \min_{\mu \in \Pi_{d,\epsilon}^{f \circ A}(x_0)} \mu [(f \circ A)'_{\epsilon}(x_0; d) - (f \circ A)'_{\epsilon}(x_0; d)] \\ &= \min_{\mu \in \Pi_{Ad,\epsilon}^f(Ax_0)} \mu [f'_{\epsilon}(Ax_0; Ad) - f'_{\epsilon}(Ax_0; Ad)] \\ &= f''_{\epsilon}(Ax_0; Ad, Ad). \end{aligned}$$

□.

Dans ce chapitre, nous avons mis en évidence la semi-continuité inférieure de  $f_\varepsilon''(x; d, d)$  en fonction de  $x$ , la non-continuité et non-convexité en fonction de  $d$ .

Nous avons également signalé des pistes de recherches sur des règles de calcul.

Nous insistons sur le fait que ce chapitre est loin de clore l'étude de  $f_\varepsilon''$  comme fonction du point  $x$  et de la direction  $d$ .

---

ANNEXES .



## Annexe 1

1. La fonction  $q_{d,\varepsilon}(\cdot) : d \mapsto \frac{f(x_0+d) - f(x_0) + \varepsilon}{d}$  est semi-continue inférieurement.

Prouvons que  $q_{d,\varepsilon}(d)$  est s.c.i. en vérifiant que, pour chaque  $\alpha \in \mathbb{R}$ , l'ensemble  $S = \{d \mid q_{d,\varepsilon}(d) \leq \alpha\}$  est fermé.

L'ensemble  $S$  est encore  $\{d \mid f(x_0+d) - f(x_0) + \varepsilon - \alpha d \leq 0\}$ .

Considérons une suite  $(d_m)_m$  de points de  $S$  convergent vers  $d$  et montrons que  $d$  est un élément de  $S$ .

$d_m \in S$  signifie que  $f(x_0+d_m) - f(x_0) + \varepsilon - \alpha d_m \leq 0$ .

Faisons tendre  $m$  vers l'infini. Comme  $f$  est continue, nous obtenons :

$$f(x_0+d) - f(x_0) + \varepsilon - \alpha d \leq 0$$

ce qui est équivalent à  $d \in S$ .

□.

2. La fonction  $q_{d,\varepsilon}(\cdot)$  est quasi convexe.

Prouvons que quelque soient  $d_1$  et  $d_2 \in \mathbb{R}_0^+$  et  $0 < \mu < 1$

$$q_{d,\varepsilon}(\mu d_1 + (1-\mu)d_2) \leq \max\{q_{d,\varepsilon}(d_1), q_{d,\varepsilon}(d_2)\}.$$

Sans perdre de généralité, supposons que  $q_{d,\varepsilon}(d_1) \geq q_{d,\varepsilon}(d_2)$ ,

$$\text{i.e. } [f(x_0+d_2) - f(x_0) + \varepsilon] \leq \frac{d_2}{d_1} [f(x_0+d_1) - f(x_0) + \varepsilon]$$

$$\begin{aligned}
 q_{d,\varepsilon}(\mu d_1 + (1-\mu)d_2) &= \frac{f(x_0 + (\mu d_1 + (1-\mu)d_2)) - f(x_0) + \varepsilon}{\mu d_1 + (1-\mu)d_2} \\
 &\leq \frac{\mu f(x_0 + d_1) + (1-\mu)f(x_0 + d_2) - f(x_0) + \varepsilon}{\mu d_1 + (1-\mu)d_2}
 \end{aligned}$$

car  $f_d$  est convexe.

$$\text{or } f(x_0) = \mu f(x_0) + (1-\mu)f(x_0)$$

$$\text{et } \varepsilon = \mu \varepsilon + (1-\mu)\varepsilon$$

et donc

$$\begin{aligned}
 q_{d,\varepsilon}(\mu d_1 + (1-\mu)d_2) &\leq \frac{\mu f(x_0 + d_1) - f(x_0) + \varepsilon}{\mu d_1 + (1-\mu)d_2} + (1-\mu) \frac{f(x_0 + d_2) - f(x_0) + \varepsilon}{\mu d_1 + (1-\mu)d_2} \\
 &\leq \frac{\mu f(x_0 + d_1) - f(x_0) + \varepsilon}{[\mu d_1 + (1-\mu)d_2]} + (1-\mu) \frac{d_2}{d_1} \frac{f(x_0 + d_1) - f(x_0) + \varepsilon}{\mu d_1 + (1-\mu)d_2} \\
 &= \frac{f(x_0 + d_1) - f(x_0) + \varepsilon}{\mu d_1 + (1-\mu)d_2} \cdot \frac{\mu d_1 + (1-\mu)d_2}{d_1} \\
 &= q_{d,\varepsilon}(d_1) \\
 &= \max\{q_{d,\varepsilon}(d_1), q_{d,\varepsilon}(d_2)\}.
 \end{aligned}$$

□

3. La fonction  $q_{d,\varepsilon}(\cdot)$  est localement lipschitz.

Pour cela, nous montrons que :

$$\forall x \in \mathbb{R} \quad \exists B(x, \delta), \exists \eta > 0 : y, z \in B(0, \delta) \Rightarrow |f(y) - f(z)| < \eta \|y - z\|$$

Fixons  $d \in \mathbb{R}$

$$\delta = d/2.$$

$$\mu_1, \mu_2 \in B(d, d/2)$$

$$\begin{aligned} |q_{d,\varepsilon}(\mu_1) - q_{d,\varepsilon}(\mu_2)| &= \left| \frac{f(x_0 + \mu_1 d) - f(x_0) + \varepsilon}{\mu_1} - \frac{f(x_0 + \mu_2 d) - f(x_0) + \varepsilon}{\mu_2} \right| \\ &= \left| \frac{\mu_2 f(x_0 + \mu_1 d) - \mu_1 f(x_0 + \mu_2 d) - \mu_2 f(x_0) + \mu_1 f(x_0) + (\mu_1 - \mu_2)\varepsilon}{\mu_1 \mu_2} \right| \end{aligned}$$

Or  $\mu_1$  et  $\mu_2 > d/2$  et donc  $\mu_1 \cdot \mu_2 > \frac{d^2}{4}$ .

$$|q_{d,\varepsilon}(\mu_1) - q_{d,\varepsilon}(\mu_2)| \leq \frac{4}{d^2} \left[ |\mu_2 - \mu_1| \varepsilon + |\mu_1 - \mu_2| f(x_0) + |\mu_2 f(x_0 + \mu_1 d) - \mu_1 f(x_0 + \mu_2 d)| + |(\mu_2 - \mu_1) f(x_0 + \mu_2 d)| \right]$$

$$\leq \frac{4}{d^2} \left[ |\mu_2 - \mu_1| \varepsilon + |\mu_1 - \mu_2| f(x_0) + \frac{1}{2} K |\mu_1 - \mu_2| \|d\| + K' |\mu_2 - \mu_1| \right] \quad \text{car } f \text{ est bornée}$$

sur un voisinage de  $x_0$ .

$$\leq \frac{4}{d^2} \left[ \varepsilon + f(x_0) + \frac{3d}{2} K \|d\| + K' \right] |\mu_2 - \mu_1|$$

Dès lors :

$$|q_{d,\varepsilon}(\mu_1) - q_{d,\varepsilon}(\mu_2)| \leq \Pi |\mu_2 - \mu_1| \quad \text{et}$$

$q_{d,\varepsilon}(\cdot)$  est localement lipschitz.

□

## Annexe 2.

### Règles de calculs sur $f'_\varepsilon(x_0; d)$

Puisque la fonction  $d \mapsto f'_\varepsilon(x_0; d)$  est la fonction d'appui de l'ensemble  $\mathcal{D}_\varepsilon f(x_0)$ , les règles de calculs sur la dérivée directionnelle à  $\varepsilon$  près découlent de celles sur le  $\varepsilon$ -sous-différentiel et les fonctions d'appui. Énonçons les résultats principaux dont les détails se trouvent en ([1], p.11-15)

• Dérivée directionnelle à  $\varepsilon$  près de  $f_1 + f_2$ , de  $f \circ A$ .

---

— Étant donné deux fonctions  $f_1$  et  $f_2$ , on a

$$(f_1 + f_2)'_\varepsilon(x_0; d) = \max_{\substack{\varepsilon_1 \geq 0, \varepsilon_2 \geq 0 \\ \varepsilon_1 + \varepsilon_2 = \varepsilon}} \{ f'_{\varepsilon_1}(x_0; d) + f'_{\varepsilon_2}(x_0; d) \}$$

— Si  $A$  est linéaire de  $\mathbb{R}^n$  dans  $\mathbb{R}^m$  et  $f$  convexe de  $\mathbb{R}^m$  dans  $\mathbb{R}$ , alors

$$(f \circ A)'_\varepsilon(x_0; d) = f'_\varepsilon(Ax_0; Ad).$$

• Dérivée directionnelle d'un maximum de fonctions convexes  $f_i$  ( $i=1 \dots m$ ).

---

Soit  $f = \max_{i=1 \dots m} f_i$

$$f'_\varepsilon(x_0; d) = \max_{\alpha_i, \varepsilon_i} \left\{ \sum_{i=1}^m (\alpha_i f_i)'_{\varepsilon_i}(x_0; d) \right\}$$

tg  $\forall i=1 \dots m \quad \alpha_i \geq 0, \varepsilon_i \geq 0, \sum_{i=1}^m \alpha_i = 1$  et

$$\sum_{i=1}^m \varepsilon_i + f(x_0) - \sum_{i=1}^m \alpha_i f_i(x_0) = \varepsilon$$



### Annexe 3.

Soit  $\lambda > 0$  tel que  $q_{d,\varepsilon}(\lambda) \leq f'_\varepsilon(x_0; d) + \alpha$ .

Montrons que  $\lambda \geq \frac{\varepsilon}{2\pi \|d\| + \alpha}$

---

Soit  $\lambda > 0$  vérifiant  $q_{d,\varepsilon}(\lambda) \leq f'_\varepsilon(x_0; d) + \alpha$ .

Alors, pour tout  $\mu > 0$

$$\frac{f(x_0 + \lambda d) - f(x_0) + \varepsilon}{\lambda} \leq \frac{f(x_0 + \mu d) - f(x_0) + \varepsilon}{\mu} + \alpha$$

Ainsi

$$\frac{\varepsilon}{\lambda} \leq \frac{f(x_0 + \mu d) - f(x_0)}{\mu} - \frac{f(x_0 + \lambda d) - f(x_0)}{\lambda} + \alpha + \frac{\varepsilon}{\mu}$$

et puisque  $f$  est L.L. :

$$\frac{\varepsilon}{\lambda} \leq 2\pi \|d\| + \alpha + \frac{\varepsilon}{\mu}$$

Puisque cette inégalité est vraie pour chaque  $\mu$ , on a

$$\lambda \geq \frac{\varepsilon}{2\pi \|d\| + \alpha}$$

□

---

# Annexe 4.

$$\partial_{\varepsilon} f(x_0)_d = \partial \psi_{\partial_{\varepsilon} f(x_0)}^*(d)$$

$$\text{Gi } \psi_{\partial_{\varepsilon} f(x_0)}^*(d) = f'_{\varepsilon}(x_0; d) .$$

Rappelons que

$$\partial_{\varepsilon} f(x_0)_d = \{x^* \in \partial_{\varepsilon} f(x_0) \mid \langle x^*, d \rangle = f'_{\varepsilon}(x_0; d)\}$$

$$\partial \psi_{\partial_{\varepsilon} f(x_0)}^*(d) = \{x^* \mid \forall d': \psi_{\partial_{\varepsilon} f(x_0)}^*(d') \geq \psi_{\partial_{\varepsilon} f(x_0)}^*(d) + \langle x^*, d' - d \rangle\}$$

Prouvons l'égalité de ces deux ensembles par la méthode de la double inclusion.

1). Soit  $x^* \in \partial_{\varepsilon} f(x_0)_d$  i.e.  $x^* \in \partial_{\varepsilon} f(x_0)$  et  $\langle x^*, d \rangle = f'_{\varepsilon}(x_0; d)$  (1)

quelque soit  $d'$ , nous aurons toujours

$$\begin{aligned} f'_{\varepsilon}(x_0; d') &\geq \langle x^*, d' \rangle \\ &\geq \langle x^*, d \rangle + \langle x^*, d' - d \rangle \\ &\geq f'_{\varepsilon}(x_0; d) + \langle x^*, d' - d \rangle \quad \text{par (1)} \end{aligned}$$

et donc  $x^* \in \partial \psi_{\partial_{\varepsilon} f(x_0)}^*(d)$

2) soit  $x^* \in \partial \psi_{\partial_{\varepsilon} f(x_0)}^*(d)$  i.e.  $\forall d' f'_{\varepsilon}(x_0; d') \geq f'_{\varepsilon}(x_0; d) + \langle x^*, d' - d \rangle$  (2)

Vérifions que  $\langle x^*, d \rangle = f'_{\varepsilon}(x_0; d)$

Choisissons  $d' = 0$  et  $d' = d$  dans (2). Nous obtenons

successivement :

$$f'_\varepsilon(x_0; d) \leq \langle x^*, d \rangle$$

$$f'_\varepsilon(x_0; d) \geq \langle x^*, d \rangle.$$

$$\text{On a donc } f'_\varepsilon(x_0; d) = \langle x^*, d \rangle.$$

Si nous combinons cette égalité avec (e), il découle que  
 $\forall d', f'_\varepsilon(x_0; d') \geq \langle x^*, d' \rangle$  i.e.  $x^* \in \partial_\varepsilon f(x_0)$ .

□

### Annexe 5.

1. La fonction  $\lambda \mapsto f'(x_0 + \lambda d; d)$  est monotone.

On sait que  $\partial f(\cdot)$  est monotone. Ceci se traduit par :

$\forall x_1, x_2 \in \mathbb{R}^n, \forall y_1 \in \partial f(x_1)$  et  $\forall y_2 \in \partial f(x_2)$  :

$$\langle x_1 - x_2, y_1 - y_2 \rangle \geq 0.$$

(1)

Prenons  $\lambda_1 < \lambda_2$ ,  $x_1 = x_0 + \lambda_1 d$

$$x_2 = x_0 + \lambda_2 d.$$

(1) se réécrit :

$$(\lambda_1 - \lambda_2) \langle d, y_1 - y_2 \rangle \geq 0.$$

Comme  $(\lambda_1 - \lambda_2) < 0$ , on a  $\langle d, y_1 - y_2 \rangle \leq 0$ .

et donc  $\langle d, y_1 \rangle \leq \langle d, y_2 \rangle \quad \forall y_1 \in \partial f(x_1) \text{ et } \forall y_2 \in \partial f(x_2)$  (2)

d'où  $\sup_{\partial f(x_1)} \langle d, y_1 \rangle \leq \sup_{\partial f(x_2)} \langle d, y_2 \rangle$

On envoie  $f'(x_0 + \lambda_1 d; d) \leq f'(x_0 + \lambda_2 d; d)$

□

2. Prouvons que :

$$D''_x f(x_0; d) = \lim_{d \rightarrow 0^+} \frac{1}{d} [\langle \partial f(x_0 + d); d \rangle - f'(x_0; d)].$$


---

Par (2), on a :  $f'(x_0 + d_1; d) \leq \langle d, y_1 \rangle \quad \forall y_1 \in \partial f(x_0 + d_1; d)$   
 $\frac{f'(x_0 + d_1; d) - f'(x_0; d)}{d} \leq \frac{\langle d, y_1 \rangle - f'(x_0; d)}{d} \quad \forall y_1 \in \partial f(x_0 + d_1; d)$

Prenons  $d_1 < d_2 < d$  et passons à la limite dans l'inégalité

ci-dessus :

$$\lim_{d \rightarrow 0} \frac{f'(x_0 + d_1; d) - f'(x_0; d)}{d} \leq \lim_{d \rightarrow 0} \frac{\langle d, y_1 \rangle - f'(x_0; d)}{d}$$

$$\forall y_1 \in \partial f(x_0 + d_1; d).$$

$$\text{Et } f'(x_0 + d_2; d) \geq \langle y_2, d \rangle \quad \forall y_2 \in \partial f(x_0 + d_2; d)$$

et donc

$$\lim_{d \rightarrow 0^+} \frac{f'(x_0 + d_2; d) - f'(x_0; d)}{d} \geq \lim_{d \rightarrow 0^+} \frac{\langle y_2, d \rangle - f'(x_0; d)}{d}$$

D'où

$$\lim_{d \rightarrow 0^+} \frac{f'(x_0 + d; d) - f'(x_0; d)}{d} = \lim_{d \rightarrow 0^+} \frac{1}{d} [\langle \partial f(x_0 + d); d \rangle - f'(x_0; d)]$$

□



## *BIBLIOGRAPHIE.*

## BIBLIOGRAPHIE

- (1). J-B. HIRIART-URRUTY,  $\epsilon$ -subdifferential calculus.  
(février 1980)
- (2). J-B. HIRIART-URRUTY, Approximating a second-order directional derivative for nonsmooth convex functions.
- (3). J-B. HIRIART-URRUTY, Limiting behaviour of the approximate first-order and second-order directional derivatives for a convex function. (juin 81).
- (4). J-B. HIRIART-URRUTY, The approximate first-order and second-order directional derivatives for a convex function.
- (5). A. AUSLENDER, On the differential properties of the support of the  $\epsilon$ -subdifferential of a convex function.  
(1981)
- (6). D.P. BERTSEKAS and S.K. MITTER, A descent numerical method for optimization problems with nondifferentiable cost functionals. (1973). *Siam J. Control.* Vol. 11. No.4 November 1973. P. 637.
- (7). R.T. ROCKAFELLAR, *Convex Analysis*.
- (8). I. EKELAND et R. TEMAN, *Analyse convexe et problèmes variationnels*, Bordas. Paris (1958).
- (9). H. BUSEMAN, *Convex surfaces*, (1958) P. 10.
- (10). R.T. ROCKAFELLAR, Monotone operators and the proximal point algorithm. *Siam J. Control optimization* 14 (1976), P. 877-898.
- (11). F. MIGNOT, *J. Functional analysis*, 22, 1976, P.130-185.
- (12). V.F. DEM'YANOV and V.N. MALOZEMOV, *Introduction to minimax*, John Wiley and Sons, (1974).
- (13). E. ASPLUND and R.T. ROCKAFELLAR, Gradients of convex functions, *Trans. Amer. Math. Soc.* 139 (1969), P.443-467.
- (14). J-B. HIRIART-URRUTY, Lipschitz  $r$ -continuity of the approximate subdifferential of a convex function (1980).